

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 10:29:22 : Search time 168.5 seconds
(without alignments)
1912.600 Million cell updates/sec

Title: US-09-980-265-9
Perfect score: 20
Sequence: 1 ttagtcttagttagt 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 60155 seqs, 8056825618 residues

Total number of hits satisfying chosen parameters: 120310

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl_HTG:*

1: gb_htg:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	19	95.0	186542	1	AL844197 Danio rer
2	18.4	92.0	5421	1	AC019958 Drosophi
3	18.4	92.0	38069	1	AC017255 Drosophi
4	18.4	92.0	80679	1	AC103576 Rattus no
5	18.4	92.0	106651	1	AC106466 Rattus no
6	18.4	92.0	155877	1	AC128449 Rattus no
7	18.4	92.0	176400	1	AC113369 Homo sapi
8	18.4	92.0	178970	1	AC012589 Homo sapi
9	18.4	92.0	187431	1	AC023241 Homo sapi
10	18.4	92.0	213308	1	AC079164 Mus muscu
11	18.4	92.0	217953	1	AC122334 Mus muscu
12	18.4	92.0	241585	1	AL844536 Mus muscu
13	18.4	87.0	15997	1	AC015120 Drosophi
14	17.4	87.0	102832	1	AL713963 Danio rer
15	17.4	87.0	147734	1	AL713862 Danio rer
16	17.4	87.0	148208	1	AC010951 Homo sapi
17	17.4	87.0	153328	1	AC122111 Rattus no
18	17.4	87.0	159295	1	CNS079P3 Rattus no
19	17.4	87.0	167683	1	AC079251 Homo sapi
20	17.4	87.0	169179	1	AC118888 Rattus no
21	17.4	87.0	172006	1	AC110979 Rattus no
22	17.4	87.0	173471	1	AC101841 Mus muscu
23	17.4	87.0	183450	1	AC119322 Rattus no
24	17.4	87.0	186581	1	AC130415 Homo sapi
25	17.4	87.0	195075	1	AC122291 Mus muscu
26	17.4	87.0	195444	1	AC107399 Homo sapi
27	17.4	87.0	209572	1	AC099615 Mus muscu
28	17.4	87.0	219491	1	AC092256 Mus muscu
29	17.4	87.0	225410	1	AC113592 Mus muscu
30	17.4	87.0	227884	1	AC114995 Mus muscu
31	17.4	87.0	247102	1	AC123854 Mus muscu
32	17.4	87.0	263211	1	AC098458 Rattus no

C 33	17.4	87.0	263546	1	AC099416 Mus muscu
C 34	17.4	87.0	265861	1	AC069274 Mus muscu
C 35	17.4	87.0	267052	1	AC122872 Mus muscu
C 36	17.4	87.0	276082	1	AC112365 Rattus no
C 37	16.8	84.0	8659	1	AC117270 Dictyoste
C 38	16.8	84.0	14921	1	AC015356 Drosophi
C 39	16.8	84.0	15707	1	AC019964 Drosophi
40	16.8	84.0	15792	1	AC014334 Drosophi
41	16.8	84.0	35974	1	AC015005 Drosophi
42	16.8	84.0	44765	1	AC116467 Mus muscu
43	16.8	84.0	50100	1	AC101025 Mus muscu
C 44	16.8	84.0	52486	1	AC101922 Mus muscu
C 45	16.8	84.0	55036	1	AC100584 Mus muscu

ALIGNMENTS

RESULT 1
LOCUS AL844197/c
DEFINITION Danio rerio clone DKEX-15619, *** SEQUENCING IN PROGRESS ***, 7
unorderd pieces.
ACCESSION AL844197.2 GI:21953176
VERSION AL844197.2
KEYWORDS HTG: HTGS_PHASE1, HTGS_DRAFT, HTGS_FULLTOP.
SOURCE Danio rerio.
ORGANISM Danio rerio.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 186542)

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Submitted (18-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zface@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Jul 24, 2002 this sequence version replaced gi:21912371.

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: zface@sanger.ac.uk

----- Project Information

Center project name: zki5619

----- Summary Statistics

Assembly program: XGAP4; Version 4.5

Chemistry: Dye-terminator; 100% of reads

Consensus quality: 184125 bases at least Q40

Consensus quality: 184667 bases at least Q30

Consensus quality: 185376 bases at least Q20

Insert size: 185942; sum-of-ctrls

Insert size: 189146; 2.5% error; agarose-1p

Quality coverage: 7.02x in Q20 bases; sum-of-ctrls Quality

coverage: 6.90x in Q20 bases; agarose-1p

* NOTE: This is a 'working draft' sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 82061: contig of 82061 bp in length
* 82062 82161: gap of 100 bp
* 82162 92135: contig of 9974 bp in length
* 92136 92235: gap of 100 bp
* 92236 123057: contig of 30822 bp in length
* 123058 123157: gap of 100 bp
* 123158 133138: contig of 9961 bp in length
* 133139 133238: gap of 100 bp
* 133239 138160: contig of 4922 bp in length

FEATURES	SOURCE
18161 138260: gap of 100 bp	*
138261 175279: contig of 37019 bp in length	*
175280 175379: gap of 100 bp	*
175380 186542: contig of 11163 bp in length	*
Location/Qualifiers	
1. 186542	

BASE COUNT	61324	a	34425	c	33413	g	56780	t	600	others
ORIGIN										

AC019958/c	LOCUS	DEFINITION	5421 bp	DNA	linear	HTG 03-JAN-2000
AC019958		Drosophila melanogaster, ***	SEQUENCING IN PROGRESS	***	in order	reorder

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1	(bases 1 to 5421)	Adams, M. and Venter, J.C.	Direct Submission	
		Submitted (30-DEC-1999)	Celera Genomics, 45 West Gate Drive, Rockville, MD, USA	
		This sequence was identified as CDM-10211530 by the submitter		
		For more information on this record e-mail to fly@celera.com		
		* NOTE: This is a 'working draft' sequence.		
		* This sequence will be replaced		
		* by the finished sequence as soon as it is available and		
		* the accession number will be preserved.		

BASE COUNT	1558 a	1150 c	1146 g	1567 t
ORIGIN				

Query Match	92.0%;	Score 18.4;	DB 1;	Length 5421;
Best Local Similarity	95.0%;	Pred. No. 14;		
Matches 19;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;

RESULT 3	AC017255	LOCUS	DEFINITION	AC017255	38069 bp	DNA	linear	HTG 09-DEC-1999
ACCESSION	AC017255		Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***					
VERSION	AC017255.1		pieces.					
KEYWORDS	HTG: HTGS-PHASE2.							
SOURCE	Drosophila melanogaster.							
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidae; Drosophilidae; Drosophila.							
REFERENCE	1 (bases 1 to 38069)							
AUTHORS	Adams,M. and Venter,J.C.							
TITLE	Direct Submission							
JOURNAL	Submitted (09-DEC-1999)							
COMMENT	Rockville, MD, USA							
	This sequence was identified as CDM.1021065 by the submitter.							

COMMENT
This sequence was identified as CDM:10210065 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

BASE COUNT	12432 a	7322 c	7198 g	11117 f
ORIGIN				

RESULT 4			
AC103576	80679 bp	DNA	linear
LOCUS			HTG 12-JUL-2000
DEFINITION	AC103576		
	Rattus norvegicus clone CH230-63M13, *** SEQUENCING IN PROGRESS		
	***, 47 unordered pieces.		

REFERENCE	AUTHORS
1 (pages 1 to 80679)	
Muzny, D.M., Adams, C.C.,	Adio-Oduola, B., Ali-osman, F.R., Allen, C.,
Alsbrooks, S.L., Amaralunga, H.C.,	Aze, J.R., Ayele, M., Banks, T.,
Barberia, J., Benton, J., Bimoge, K.,	Blankenburg, K., Bonini, D.,
Bouch, J., Bowie, S., Brieva, M.,	Brown, E., Brown, M., Bryant, N.P.,
Buthey, C.C., Burch, P., Buttel, C.,	Burrell, K.L., Byrd, N.C.,
Carroll, T.F., Carter, M., Cavazos, S.R.,	Chacko, J., Chave, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I.,	Christopoulos, C.,
Cleveland, C.D., Cox, C.C., Coyle, M.D.,	Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L.,	Dederich, D.A.,

BASE COUNT	1558 a	1150 c	1146 g	1567 t
ORIGIN				

Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,K., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,U., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Locado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheswari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Monabhat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokwuo,S., Oguh,M., Okunolu,G.,
Oregunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L., Quiles,S., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Schreier,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Swatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,Y., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Unpublished
2 (bases 1 to 80679)
Worley,K.C.
Direct Submission
Submitted (29-NOV-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 80679)
Worley,K.C.
Direct Submission
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:18846139.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GIGY
Center clone name: CH230-63N13
----- Summary Statistics
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 33297 bases at least Q40
Consensus quality: 35946 bases at least Q30
Consensus quality: 38552 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 47 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1171: contig of 1171 bp in length
* 1172 1271: gap of unknown length
* 2868: contig of 1597 bp in length
* 2968: gap of unknown length
* 2969 4131: contig of 1163 bp in length
* 4132 4231: gap of unknown length
* 4232 5312: contig of 1081 bp in length
* 5313 5412: gap of unknown length
* 5413 6870: contig of 1458 bp in length
* 6871 6970: gap of unknown length
* 6971 7991: contig of 1021 bp in length
* 7992 8091: gap of unknown length
* 8092 9126: contig of 1035 bp in length
* 9127 9227: gap of unknown length
* 9228 10468: contig of 1242 bp in length
* 10469 10569: gap of unknown length
* 10570 11619: contig of 1051 bp in length
* 11620 11719: gap of unknown length
* 11720 12894: contig of 1175 bp in length
* 12895 12994: gap of unknown length
* 12995 14178: contig of 1184 bp in length
* 14179 14278: gap of unknown length
* 14279 15615: contig of 1337 bp in length
* 15616 15715: gap of unknown length
* 15716 17016: contig of 1301 bp in length
* 17017 17116: gap of unknown length
* 17117 18233: contig of 1117 bp in length
* 18234 18333: gap of unknown length
* 18334 19665: contig of 1332 bp in length
* 19666 19765: gap of unknown length
* 19766 21151: contig of 1386 bp in length
* 21152 22739: contig of 1488 bp in length
* 22740 22839: gap of unknown length
* 22840 23909: contig of 1070 bp in length
* 23910 24009: gap of unknown length
* 24010 25502: contig of 1493 bp in length
* 25503 25602: gap of unknown length
* 25603 26796: contig of 1194 bp in length
* 26797 26896: gap of unknown length
* 26897 28120: contig of 1224 bp in length
* 28121 28220: gap of unknown length
* 28221 29944: contig of 1724 bp in length
* 29945 30044: gap of unknown length
* 30045 32238: contig of 2194 bp in length
* 32239 32338: gap of unknown length
* 32339 33569: contig of 1231 bp in length
* 33570 33669: gap of unknown length
* 33670 35660: contig of 1891 bp in length
* 35661 35660: gap of unknown length
* 35661 36788: contig of 1128 bp in length
* 36789 36888: gap of unknown length
* 36889 38239: contig of 1351 bp in length
* 38240 38339: gap of unknown length
* 38340 39574: gap of 1235 bp in length
* 39575 39674: gap of unknown length
* 39675 40719: contig of 1045 bp in length
* 40720 40819: gap of unknown length
* 40820 42444: contig of 1625 bp in length
* 42445 42544: gap of unknown length
* 42545 44516: contig of 1972 bp in length
* 44517 44616: gap of unknown length
* 44617 46392: contig of 1776 bp in length
* 46393 46492: gap of unknown length
* 46493 47848: contig of 1356 bp in length
* 47849 47948: gap of unknown length
* 47949 49380: contig of 1432 bp in length
* 49381 49480: gap of unknown length
* 49481 50992: contig of 1512 bp in length
* 50993 51092: gap of unknown length
* 51093 53214: contig of 2122 bp in length
* 53215 53314: gap of unknown length
* 53315 54686: contig of 1772 bp in length
* 54687 54786: gap of unknown length
* 54787 56476: contig of 1690 bp in length

Location	Contig	bp in length
55477	gap of unknown length	
55624	contig of 2048 bp in length	
55625	gap of unknown length	
55825	contig of 2587 bp in length	
61312	gap of unknown length	
61412	contig of 1996 bp in length	
63408	gap of unknown length	
63412	contig of 1905 bp in length	
63413	gap of unknown length	
63512	gap of unknown length	
65121	contig of 2609 bp in length	
68221	gap of unknown length	
71194	contig of 2973 bp in length	
71195	gap of unknown length	
71295	contig of 2360 bp in length	
73654	gap of unknown length	
73755	contig of 2705 bp in length	
76460	gap of unknown length	
76550	contig of 4120 bp in length	

[illegible]

Query Match	92.0%;	Score 18.4;	DB 1;	Length 80679;
Best Local Similarity	95.0%;	Pred. No. 17;		
Best Match	19;	Conservative	0;	Mismatches 1;
				Indels 0;
				Gaps 0

```

QY      1  TTAGTCTTAGTTATTAGTT  20
          |||||  |||||  |||||  ||
Db  41343  TTAGTTATTAGTTATTAGTT  41362

```

[illegible]

REFERENCE AUTHORS

1 (bases 10 to 106651)

Muzny, D. M., Adams, C., Adio-Oduola, B., Ali-osman, F. R., Allen, C., Alstbrooks, S. L., Amaralune, H. C., Are, J. R., Ayela, M., Banks, J., Barbieri, J., Benton, J., Bimage, K., Blankenburg, K., Bonath, D., Bouch, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N. P., Buhay, C., Burch, P., Burkett, C., Burrell, K. L., Byrd, N. C., Carrott, T. F., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R., Davila, M. L., Davis, C., Davy-Carroll, L., Dederich, D. A., Delaney, K. R., Delgado, O., Degen, A. L., Ding, Y., Dinh, H. J., Douthett, K. J., Draper, H., Dunn-Rocha, S., Durbin, K. J., Earhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homas, F., Howard, S., Huber, J., Hulys, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kover, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L. J., Li, Z., Lichtharge, O., Lieu, C., Liu, J., Liu, W., Loulsegod, H., Lozado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Mattindale, A., Martinez, E.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Massey, E., Mawmney, E., McLeod, M.P., Meador, M., Mei, G., Melker, M.,
 Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
 Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
 Nguyen, N., Nickerson, E., Nwokwenkwo, S., Ogun, M., Okunodu, G.,
 Orangunye, N., Oyedero, R., Pace, A., Payton, B., Peery, J., Perez, L.,
 Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
 Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
 Scherer, S., Scott, G., Shen, H., Shoshitali, N., Sisson, I.,
 Sodergren, E., Sonatke, T., Sparks, A., Stanley, H., Stone, H.,
 Sutton, A., Swatck, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
 Tansley, J., Taylor, C., Taylor, T., Teifrod, B., Thomas, N., Thomas, S.,
 Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, Q.,
 Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
 Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
 Wu, C., Wu, Y., Wu, Y.F., Zhou, D., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.
 Direct Submission
 2 (bases 1 to 106651)
 Unpublished
 Morley, K.C.
 Direct Submission
 Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 106651)
 Morley, K.C.
 Direct Submission
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 12, 2002 this sequence version replaced gi:18138988.

 Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu

 Project Information
 Center project name: GKXR
 Center clone name: CH230-211B18

 Summary Statistics
 Sequencing vector: Plasmid:
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap, version 0.990329
 Consensus quality: 79275 bases at least Q40
 Consensus quality: 83497 bases at least Q30
 Consensus quality: 87549 bases at least Q20

 NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html)
 * NOTE: This is a "working draft" sequence. It currently
 * consists of 45 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 1
 1130: contig of 1130 bp in length
 1131 1230: gap of unknown length
 1131 1243: contig of 1013 bp in length
 2244 2343: gap of unknown length
 2244 2344 3369: contig of 1026 bp in length
 3370 3470 5003: contig of 1534 bp in length
 3470 5103: gap of unknown length
 5104 6657: contig of 1554 bp in length
 6658 6758: gap of unknown length
 6758 7841: contig of 1083 bp in length
 7841 7941: gap of unknown length
 7941 10020: contig of 2080 bp in length
 10021 10120: gap of unknown length
 10121 11477: contig of 1357 bp in length
 11478 11577: gap of unknown length

* 11578 13365: contig of 1788 bp in length
* 13366 13465: gap of unknown length
* 13466 14853: contig of 1388 bp in length
* 14854 14953: gap of unknown length
* 14954 16290: contig of 1337 bp in length
* 16291 16390: gap of unknown length
* 16391 17509: contig of 1119 bp in length
* 17510 17609: gap of unknown length
* 17610 19000: contig of 1391 bp in length
* 19001 19100: gap of unknown length
* 19101 20944: contig of 1844 bp in length
* 20945 21044: gap of unknown length
* 21045 22651: contig of 1607 bp in length
* 22652 22751: gap of unknown length
* 22752 25058: contig of 2307 bp in length
* 25059 26156: gap of unknown length
* 26157 26817: contig of 1658 bp in length
* 26818 26916: gap of unknown length
* 26917 28641: contig of 1725 bp in length
* 28642 28741: gap of unknown length
* 28742 30799: contig of 2058 bp in length
* 30800 30899: gap of unknown length
* 30900 32351: contig of 1452 bp in length
* 32352 32451: gap of unknown length
* 32452 34522: contig of 2071 bp in length
* 34523 34622: gap of unknown length
* 34623 36509: contig of 1887 bp in length
* 36510 36609: gap of unknown length
* 36610 38251: contig of 1642 bp in length
* 38252 38351: gap of unknown length
* 38352 40309: contig of 1958 bp in length
* 40310 40409: gap of unknown length
* 40410 42792: contig of 2383 bp in length
* 42793 42892: gap of unknown length
* 42893 45780: contig of 2888 bp in length
* 45781 45880: gap of unknown length
* 45881 49221: contig of 3341 bp in length
* 49222 49321: gap of unknown length
* 49322 51468: contig of 2147 bp in length
* 51469 51568: gap of unknown length
* 51569 53749: contig of 2181 bp in length
* 53750 53849: gap of unknown length
* 53850 55748: contig of 1899 bp in length
* 55749 55848: gap of unknown length
* 55849 59788: contig of 3940 bp in length
* 59789 59888: gap of unknown length
* 59889 62649: contig of 2761 bp in length
* 62650 62749: gap of unknown length
* 62750 65930: contig of 3181 bp in length
* 65931 66030: gap of unknown length
* 66031 68484: contig of 2454 bp in length
* 68485 68584: gap of unknown length
* 68585 70509: contig of 1925 bp in length
* 70510 70609: gap of unknown length
* 70610 73606: contig of 2997 bp in length
* 73607 73706: gap of unknown length
* 73707 76448: contig of 2742 bp in length
* 76449 76548: gap of unknown length
* 76549 79392: contig of 2844 bp in length
* 79393 79492: gap of unknown length
* 79493 83319: contig of 2827 bp in length
* 83320 82419: gap of unknown length
* 82420 85792: contig of 3373 bp in length
* 85793 85892: gap of unknown length
* 85893 88433: contig of 2541 bp in length
* 88434 88533: gap of unknown length
* 88534 92245: contig of 3712 bp in length
* 92246 92345: gap of unknown length
* 92346 96997: contig of 4652 bp in length
* 96998 97097: gap of unknown length
* 97098 102010: contig of 4913 bp in length
* 102011 102110: gap of unknown length
* 102111 106651: contig of 4541 bp in length.

FEATURES
source Location/Qualifiers
1..106651
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-211B18"
BASE COUNT 29303 a 19724 c 19632 g 30261 t 7731 others
ORIGIN
Query Match 92.0%; Score 18.4; DB 1; Length 106651;
Best Local Similarity 95.0%; Pred. No. 17;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTTCCTAGTATTACTT 20
DB 83573 TTAGTATTAGTATTACTT 83554
RESULT 6
AC128449 155877 bp DNA linear HTG 19-JUL-2002
LOCUS Rattus norvegicus clone CH230-117D11, *** SEQUENCING IN PROGRESS
DEFINITION *** 85 unordered pieces.
ACCESSION AC128449
VERSION AC128449.1 GI:21909143
KEYWORDS HTG: HTGS_PHASE1.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 155877)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooks,S.L., Amaralunge,H.C., Are,J.R., Ayete,M., Banks,T.,
Barbata,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,U., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Devila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.,
Douthwaite,K.J., Draper,H., Dugan-Kocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Honsl,F., Howard,S., Huber,J., Hulik,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlssoon,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,M., Louisgeed,H.,
Lorado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokkenko,S., Ogun,M., Okunnu,G.,
Oreagunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojucokan,I., Rolfe,M., Ruiz,S., Savary,G.,
Scherer,S., Scott,G., Shen,H., Shooshitari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Swalek,A., Tabor,P., Tameis,A., Tameis,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczek,R., Woodden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
TITLE Direct Submission
JOURNAL Unpublished

REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 155877)
Worley, K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

----- Genome Center -----

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information -----

Center project name: GTWL

Center clone name: CH230-117D11

----- Summary Statistics -----

Sequencing vector: Plasmid;

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 73163 bases at least Q40

Consensus quality: 82138 bases at least Q30

Consensus quality: 87186 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently
* consists of 85 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1062: contig of 1062 bp in length
* 1063 1162: gap of unknown length
* 1163 2171: contig of 1009 bp in length
* 2172 2271: gap of unknown length
* 2272 3321: contig of 1050 bp in length
* 3322 3421: gap of unknown length
* 3422 4463: contig of 1042 bp in length
* 4464 4563: gap of unknown length
* 4564 5587: contig of 1024 bp in length
* 5588 6712: gap of unknown length
* 6713 6812: gap of unknown length
* 6813 7822: contig of 1010 bp in length
* 7823 7922: gap of unknown length
* 7923 9753: contig of 1831 bp in length
* 9754 9853: gap of unknown length
* 9854 10887: contig of 1034 bp in length
* 10888 10987: gap of unknown length
* 10989 12401: contig of 1414 bp in length
* 12402 12501: gap of unknown length
* 12502 14037: contig of 1536 bp in length
* 14038 14137: gap of unknown length
* 14139 15197: contig of 1060 bp in length
* 15198 15297: gap of unknown length
* 15299 16644: contig of 1347 bp in length
* 16645 16744: gap of unknown length
* 16745 18177: contig of 1433 bp in length
* 18178 18277: gap of unknown length
* 18279 19658: contig of 1381 bp in length
* 19659 19758: gap of unknown length
* 19759 20806: contig of 1048 bp in length
* 20807 20906: gap of unknown length
* 20907 22312: contig of 1406 bp in length
* 22313 22412: gap of unknown length
* 22413 23922: contig of 1510 bp in length
* 23923 24022: gap of unknown length
* 24023 25433: contig of 1411 bp in length
* 25434 25533: gap of unknown length
* 25534 26809: contig of 1276 bp in length
* 26810 26910: gap of unknown length
* 26910 28404: contig of 1495 bp in length

28405 28504: gap of unknown length
* 28505 29670: contig of 1166 bp in length
* 29671 29770: gap of unknown length
* 29771 30994: contig of 1224 bp in length
* 30995 31094: gap of unknown length
* 31095 32675: contig of 1561 bp in length
* 32676 32775: gap of unknown length
* 32776 34130: contig of 1355 bp in length
* 34131 34230: gap of unknown length
* 34231 35489: contig of 1259 bp in length
* 35490 35589: gap of unknown length
* 35590 37026: contig of 1437 bp in length
* 37027 37126: gap of unknown length
* 37127 38796: contig of 1670 bp in length
* 38797 38896: gap of unknown length
* 38897 40001: contig of 1105 bp in length
* 40002 40101: gap of unknown length
* 40102 41115: contig of 1014 bp in length
* 41116 41215: gap of unknown length
* 41216 42569: contig of 1354 bp in length
* 42570 42669: gap of unknown length
* 42670 43830: contig of 1161 bp in length
* 43831 43930: gap of unknown length
* 43931 45281: contig of 1351 bp in length
* 45282 45381: gap of unknown length
* 45382 47441: contig of 2060 bp in length
* 47442 47541: gap of unknown length
* 47542 49026: contig of 1485 bp in length
* 49027 49126: gap of unknown length
* 49127 50263: contig of 1137 bp in length
* 50264 50363: gap of unknown length
* 50364 52441: contig of 2078 bp in length
* 52442 52541: gap of unknown length
* 52542 53981: contig of 1440 bp in length
* 53982 54081: gap of unknown length
* 54082 55502: contig of 1421 bp in length
* 55503 55602: gap of unknown length
* 55603 56893: contig of 1251 bp in length
* 56894 56993: gap of unknown length
* 56994 58953: contig of 1960 bp in length
* 58954 59053: gap of unknown length
* 59054 60351: contig of 1298 bp in length
* 60352 60451: gap of unknown length
* 60452 61848: contig of 1397 bp in length
* 61849 61948: gap of unknown length
* 61949 64132: contig of 2184 bp in length
* 64133 64232: gap of unknown length
* 64233 65624: contig of 1392 bp in length
* 65625 65724: gap of unknown length
* 65725 66795: contig of 1071 bp in length
* 66796 66895: gap of unknown length
* 66896 68161: contig of 1266 bp in length
* 68162 68261: gap of unknown length
* 68262 69390: contig of 1129 bp in length
* 69391 69490: gap of unknown length
* 69491 70785: contig of 1295 bp in length
* 70786 70885: gap of unknown length
* 70886 73269: contig of 2384 bp in length
* 73270 73369: gap of unknown length
* 73370 75050: contig of 1681 bp in length
* 75051 75150: gap of unknown length
* 75151 76285: contig of 1135 bp in length
* 76286 76385: gap of unknown length
* 76386 78166: contig of 1761 bp in length
* 78167 78266: gap of unknown length
* 78267 80185: contig of 1919 bp in length
* 80186 80285: gap of unknown length
* 80286 81681: contig of 1396 bp in length
* 81682 81781: gap of unknown length
* 81782 83340: contig of 1559 bp in length
* 83341 83441: gap of unknown length
* 83441 85436: contig of 1996 bp in length
* 85437 85536: gap of unknown length

Query Match 92.0%; Score 18.4; DB 1; Length 155877;
Best Local Similarity 95.0%; Pred. No. 18;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTATTACTT 20
|||||
Db 46393 TTAGTCTTACTTACTTACTT 46412

RESULT 7
AC113369 176400 bp DNA linear HTG 01-MAR-2002
LOCUS Homo sapiens chromosome 5 clone RP11-119P10, WORKING DRAFT
DEFINITION
SEQUENCE, 1 unordered piece.
AC113369
AC113369.1 GI:19033483
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVEFIN.
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Eutheria; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
JOURNAL DOE Joint Genome Institute.
REFERENCE
AUTHORS Sequencing of Human Chromosome 5
TITLE 2 (bases 1 to 176400)
JOURNAL DOE Joint Genome Institute.
REFERENCE
AUTHORS Direct Submission
TITLE Submitted (01-MAR-2002) Production Sequencing Facility, DOE Joint
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 452431
Center clone name: RPCI-11_119P10

Summary Statistics
Consensus quality: 175634 bases at least Q40
Consensus quality: 176221 bases at least Q30
Consensus quality: 176375 bases at least Q20
Estimated insert size: 182000; agarose-fp estimation
Estimated insert size: 176400; sum-of-contigs estimation
Quality coverage: 12.49 in Q20 bases; agarose-fp estimation
Quality coverage: 12.88 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

FEATURES
SOURCE
1 176400: contig of 176400 bp in length.
1.176400
Location/Qualifiers
1.176400
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-119P10"
/clone_lib="RPCI human BAC library 11"

BASE COUNT 55190 a 37584 c 34956 g 48670 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 176400;
Best Local Similarity 95.0%; Pred. No. 18;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTACTTATTACTT 20
|||||

Db 78311 TTAGTATTACTTATTACTT 78330

RESULT 8
AC012589 178970 bp DNA linear HTG 24-AUG-2002
LOCUS Homo sapiens chromosome 5 clone RP11-119P10 map 5, WORKING DRAFT
DEFINITION
SEQUENCE, 23 unordered pieces.
AC012589
AC012589.5 GI:7387335
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens.
REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Eutheria; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
JOURNAL DOE Joint Genome Institute.
REFERENCE
AUTHORS Sequencing of Human Chromosome 5
TITLE 2 (bases 1 to 178970)
JOURNAL DOE Joint Genome Institute.
REFERENCE
AUTHORS Direct Submission
TITLE Submitted (30-OCT-1999) Whitehead Institute/MIT Center for Genome
JOURNAL Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE
AUTHORS 3 (bases 1 to 178970)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Baldwin, J., Barna, N., Beckler, R., Boguslavsky, L., Boukhalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collamore, A.,
Cooke, P., DeArrellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howard, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lien, C., Locke, K., Macdonald, P., Margulis, N.,
McEwan, P., McGurt, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (30-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 178970)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,
Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,
Collins, S., Collamore, A., Cooke, P., DeArrellano, K., Dewar, K.,
Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Glnde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W.,
Iliev, I., Johnson, R., Jones, C., Karatas, A., Larocque, K.,
Lamazares, R., Landers, T., Lehoczky, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Margulis, N., Matthews, C., McCarty, M.,
McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menais, L.,
Milnova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C.,
Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, S.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnapp, R., Seaman, S.,
Severy, P., Soungue, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S.,
Theodore, J., Travers, M., Travis, N., Trigglio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 3, 2000 this sequence version replaced gi:6910762.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RW/RepeatMasker.html>

COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: I3738

```

-----
Center-clone name: 119.P.10
-----
Summary statistics
Sequencing vector: M13, M7815, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap, version 0.960731
Consensus quality: 160436 bases at least Q40
Consensus quality: 168973 bases at least Q30
Consensus quality: 173202 bases at least Q20
Insert size: 177000; agarose-fp
Insert size: 176770; sum-of-contigs
Quality coverage: 4.0 in Q20 bases; agarose-fp
Quality coverage: 4.0 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1
1011: contig of 1011 bp in length
1012 1111: gap of 100 bp
1112 2722: contig of 1611 bp in length
2723 2822: gap of 100 bp
2823 4300: contig of 1478 bp in length
4301 4400: gap of 100 bp
4401 5425: contig of 1025 bp in length
5426 5525: gap of 100 bp
5526 7534: contig of 2009 bp in length
7535 7634: gap of 100 bp
7635 9780: contig of 2146 bp in length
9781 9880: gap of 100 bp
9881 12462: contig of 2582 bp in length
12463 12562: gap of 100 bp
12563 15626: contig of 3064 bp in length
15627 15726: gap of 100 bp
15727 17635: contig of 1909 bp in length
17636 17735: gap of 100 bp
17736 18894: contig of 2159 bp in length
18895 19994: gap of 100 bp
19995 25298: contig of 5304 bp in length
25299 25398: gap of 100 bp
25399 30406: contig of 5008 bp in length
30407 30506: gap of 100 bp
30507 38666: contig of 8160 bp in length
38667 38766: gap of 100 bp
38767 45443: contig of 6683 bp in length
45450 45549: gap of 100 bp
45550 55180: contig of 9631 bp in length
55181 55280: gap of 100 bp
55281 63752: contig of 8472 bp in length
63753 63855: gap of 100 bp
63856 75295: contig of 11443 bp in length
75296 75395: gap of 100 bp
75396 87581: contig of 12188 bp in length
87584 87683: gap of 100 bp
87684 98245: contig of 10552 bp in length
98246 98345: gap of 100 bp
98346 111685: contig of 13340 bp in length
11686 111785: gap of 100 bp
11786 131002: contig of 19217 bp in length
131003 131102: gap of 100 bp
131103 155070: contig of 20968 bp in length
152071 152170: gap of 100 bp
152171 178970: contig of 26800 bp in length.

Location/Qualifiers
1. 178970
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5"
/clone="RP11-119P10"

```

```

misc_feature      /clone_lib="RPC1-11 Human Male BAC"
                  1. 1011
misc_feature      /note="assembly_fragment"
                  1112. 2722
misc_feature      /note="assembly_fragment"
                  2823. 4300
misc_feature      /note="assembly_fragment"
                  4401. 5425
misc_feature      /note="assembly_fragment"
                  5526. 7534
misc_feature      /note="assembly_fragment"
                  7635. 9780
misc_feature      /note="assembly_fragment"
                  9881. 12462
misc_feature      /note="assembly_fragment"
                  12563. 13626
misc_feature      /note="assembly_fragment"
                  15727. 17635
misc_feature      /note="assembly_fragment"
                  17736. 19894
misc_feature      /note="assembly_fragment"
                  19995. 25298
misc_feature      /note="assembly_fragment"
                  25399. 30406
misc_feature      /note="assembly_fragment"
                  30507. 38666
misc_feature      /note="assembly_fragment"
                  38767. 45349
misc_feature      /note="assembly_fragment"
                  45550. 55180
misc_feature      /note="assembly_fragment"
                  55281. 63752
misc_feature      /note="assembly_fragment"
                  63853. 75295
misc_feature      /note="assembly_fragment"
                  75396. 87583
misc_feature      /note="assembly_fragment"
                  87684. 98245
misc_feature      /note="assembly_fragment"
                  98346. 111685
misc_feature      /note="assembly_fragment"
                  111786. 131002
misc_feature      /note="assembly_fragment"
                  131103. 152070
misc_feature      /note="assembly_fragment"
                  clone_end:sp6
                  vector_side:right"
                  152171. 178970
misc_feature      /note="assembly_fragment"
                  clone_end:r7
                  vector_side:left"
BASE COUNT      54765 a 37234 c 35721 g 49050 t 2200 others
ORIGIN
Query Match      92.0%: Score 18.4; DB 1; length 178970;
Best Local Similarity 95.0%: Pred. No. 18;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTATTAGTT 20
      ||||| ||||| |||||
Db 22509 TTAGTATTAGTATTAGTT 22528

RESULT 9
AC023241/c      AC023241      187431 bp      DNA      linear      HTG 01-MAR-2000
LOCUS      Homo sapiens clone RP11-14L13, WORKING DRAFT SEQUENCE, 17 unordered
DEFINITION      pieces.
ACCESSION      AC023241
VERSION      AC023241.2 GI:7138122
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens
ORGANISM

```


Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 236896 bases at least Q40
Consensus quality: 238326 bases at least Q30
Consensus quality: 239491 bases at least Q20
Insert size: 240585; sum-of-contrigs
Quality coverage: 4.96x in Q20 bases; sum-of-contrigs Quality
Coverage: 5.80x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently
consists of 11 contrigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contrigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 26384: contrig of 26384 bp in length
26385 26484: gap of 100 bp
26485 38315: contrig of 11831 bp in length
38316 38415: gap of 100 bp
38416 41944: contrig of 3529 bp in length
41945 42044: gap of 100 bp
42045 56520: contrig of 14476 bp in length
56521 56620: gap of 100 bp
56621 77754: contrig of 21134 bp in length
77755 77854: gap of 100 bp
77855 96164: contrig of 18310 bp in length
96165 96264: gap of 100 bp
96265 101917: contrig of 5653 bp in length
101918 102017: gap of 100 bp
102018 146947: contrig of 44930 bp in length
146948 147047: gap of 100 bp
147048 187542: contrig of 40495 bp in length
187543 187642: gap of 100 bp
187643 192747: contrig of 5105 bp in length
192748 192847: gap of 100 bp
192848 241585: contrig of 48738 bp in length.
Location/Qualifiers
1. 241585

FEATURES

source
1. 241585
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="2"
/clone_lib="RPC1-23"
/clone="RP23-22A15"
1. 26384
/note="assembly_fragment:01220
fragment_chain:1"
26485. 38315
/note="assembly_fragment:01223
fragment_chain:1"
38416. 41944
/note="assembly_fragment:01396
fragment_chain:1"
42045. 56520
/note="assembly_fragment:01233
fragment_chain:1"
56621. 77754
/note="assembly_fragment:00572
fragment_chain:2"
77855. 96164
/note="assembly_fragment:01046
fragment_chain:2"
96265. 101917
/note="assembly_fragment:01569
fragment_chain:2"
102018. 146947
/note="assembly_fragment:00752
fragment_chain:2"
vector_end:77
vector_side:right"
147048. 187542

/note="assembly_fragment:00109
fragment_chain:2"
187643. 192747
/note="assembly_fragment:00452
fragment_chain:2"
192848. 241585
/note="assembly_fragment:00041
fragment_chain:2"

BASE COUNT 68487 a 53697 c 52331 g 66069 t 1001 others
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 241585;
Best Local Similarity 95.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTCTTAGTTATTAGT 20
||||| |||||||||
Db 164303 TTAGTCTTAGTTATTAGT 164322

RESULT 13
AC015120 19997 bp DNA linear HTG 16-NOV-1999
LOCUS
DEFINITION
AC015120
AC015120.1 GI:6436215
VERSION
KEYWORDS
HTG: HTGS_PHASE2.
SOURCE
ORGANISM
Drosophila melanogaster.
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 19997)
AUTHORS
Adams, M. and Venter, J.C.
TITLE
Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
JOURNAL

COMMENT
This sequence was identified as CDM:10211942 by the submitter.
For further information on this sequence e-mail to fly@celera.com.
NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
1. 19997
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

FEATURES

source
1. 19997
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
BASE COUNT 5276 a 4312 c 4203 g 6206 t
ORIGIN

Query Match 87.0%; Score 17.4; DB 1; Length 19997;
Best Local Similarity 94.7%; Pred. No. 64;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTCTTAGTTATTAGT 19
||||| |||||||||
Db 1511 TTAGTCTTAGTTATTAGT 1529

RESULT 14
AL713963 102832 bp DNA linear HTG 27-JUL-2002
LOCUS
DEFINITION
Danio rerio clone BUSM1-78P07, *** SEQUENCING IN PROGRESS ***
AL713963
AL713963.9 GI:22003166
ACCESSION
KEYWORDS
HTG: HTGS_PHASE1; HTGS_ACTIVEPIN; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE
Danio rerio
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.


```

REFERENCE
AUTHORS      1 (bases 1 to 102832)
TITLE        Skuce, C.
JOURNAL      Direct Submission
              Submitted (26-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
              Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zface@sanger.ac.uk
              Clone requests: clonerequest@sanger.ac.uk
              On Jul 29, 2002 this sequence version replaced g1:22002703.
COMMENT      ----- Genome Center
              Center: Wellcome Trust Sanger Institute
              Center code: SC
              Web site: http://www.sanger.ac.uk
              Contact: zface@sanger.ac.uk
              ----- Project Information
              Center project name: dz78P07
              ----- Summary Statistics
              Assembly program: XGAP4; version 4.5
              Sequencing vector: plasmid; L08752; 100% of reads
              Chemistry: Dye-terminator; 12% of reads
              Chemistry: Dye-terminator Big Dye; 87% of reads
              Consensus quality: 102547 bases at least Q40
              Consensus quality: 102595 bases at least Q30
              Consensus quality: 102618 bases at least Q20
              Insert size: 102632; sum-of-contigs
              Insert size: 112120; 3.5% error; agarose-fp
              Quality coverage: 8.47x in Q20 bases; sum-of-contigs Quality
              coverage: 7.98x in Q20 bases; agarose-fp
              -----
              * NOTE: This is a 'working draft' sequence. It currently
              * consists of 3 contigs. The true order of the pieces
              * is not known and their order in this sequence record is
              * arbitrary. Gaps between the contigs are represented as
              * runs of N, but the exact sizes of the gaps are unknown.
              * This record will be updated with the finished sequence
              * as soon as it is available and the accession number will
              * be preserved.
              *
              * 1 19949: contig of 19949 bp in length
              * 19950 20049: gap of 100 bp
              * 20050 22338: contig of 3190 bp in length
              * 23240 23339: gap of 100 bp
              * 23340 102832: contig of 79493 bp in length.
              *
              Location/Qualifiers
                1..102832
                  /organism="Danio rerio"
                  /db_xref="taxon:7955"
                  /clone="BUSM1-78P07"
                  /clone_1id="BUSM1"
                  1..19949
                    /note="assembly-fragment:00162
                    fragment_chain:1
                    clone_end:17
                    vector_side:left"
                    20050..23239
                      /note="assembly-fragment:01428
                      fragment_chain:1"
                      23340..102832
                        /note="assembly-fragment:01396
                        fragment_chain:1"
BASE COUNT   32038 a 19328 c 19633 g 31633 t      200 others
ORIGIN
Query Match      87.0%; Score 17.4; DB 1; Length 102832;
Best Local Similarity 94.7%; Pred. No. 72;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      2 TAGTCTTACTATTACTT 20
          ||||||| |||||||
Db 77546 TAGTCTTATTATTAGTT 77564

RESULT 15
LOCUS      AL713862      147734 bp      DNA      linear      HTG 10-JUL-2002
DEFINITION Danio rerio clone BUSM1-66H22. *** SEQUENCING IN PROGRESS ***

```

```

unorderd pieces.
ACCESSION      AL713862
VERSION        AL713862.6  GI:20520522
KEYWORDS       HTG; HTCS_PHASE1; HTCS-CANCELLED.
SOURCE         zebrafish.
ORGANISM       Danio rerio
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS        Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
TITLE          Cypriniformes; Cyprinidae; Danio.
JOURNAL        1 (bases 1 to 147734)
               Babbage,A.
               Direct Submission
               Submitted (09-JUL-2002) Wellcome Trust Sanger Institute, Hinxton,
               Cambridgeshire, CB10 1SA, UK. E-mail enquiries: zface@sanger.ac.uk
               Clone requests: clonerequest@sanger.ac.uk
               On May 9, 2002 this sequence version replaced gi:20218673.
COMMENT        ----- Genome Center
               Center: Wellcome Trust Sanger Institute
               Center code: SC
               Web site: http://www.sanger.ac.uk
               Contact: zface@sanger.ac.uk
               ----- Project Information
               Center project name: dz68H22
               ----- Summary Statistics
               Assembly program: XGAP4, version 4.5
               Sequencing vector: plasmid; L08752; 100% of reads
               Chemistry: Dye-terminator; 6% of reads
               Chemistry: Dye-terminator Big Dye; 93% of reads
               Consensus quality: 140645 bases at least Q40
               Consensus quality: 142139 bases at least Q40
               Consensus quality: 143745 bases at least Q20
               Insert size: 146734; sum-of-ctrls
               Insert size: 105547; 8.3% error; agarose-fp
               Quality coverage: 15.20x in Q20 bases; sum-of-ctrls Quality
               coverage: 26.23x in Q20 bases; agarose-fp
               -----
               * NOTE: This is a 'working draft' sequence. It currently
               * consists of 11 contigs. The true order of the pieces
               * is not known and their order in this sequence record is
               * arbitrary. Gaps between the contigs are represented as
               * runs of N, but the exact sizes of the gaps are unknown.
               * This record will be updated with the finished sequence
               * as soon as it is available and the accession number will
               * be preserved.
               1
               1 2540: contig of 2540 bp in length
               * 2541 2640: gap of 100 bp
               * 2641 85546: contig of 82906 bp in length
               * 85547 85646: gap of 100 bp
               * 85647 100008: contig of 14362 bp in length
               * 100009 100108: gap of 100 bp
               * 100109 106574: contig of 6466 bp in length
               * 106575 106674: gap of 100 bp
               * 106675 126703: contig of 20029 bp in length
               * 126704 126803: gap of 100 bp
               * 126804 130027: contig of 3224 bp in length
               * 130028 130127: gap of 100 bp
               * 130128 133270: contig of 3143 bp in length
               * 133271 133370: gap of 100 bp
               * 133371 135771: contig of 2401 bp in length
               * 135772 135871: gap of 100 bp
               * 135872 138584: contig of 2713 bp in length
               * 138585 138684: gap of 100 bp
               * 138685 140983: contig of 2299 bp in length
               * 140984 141083: gap of 100 bp
               * 141084 147734: contig of 6651 bp in length.
               Location/Qualifiers
               1. 147734
               /organism="Danio rerio"
               /db_xref="taxon:7955"
               /clone="BUSM1-68H22"
               /clone_lib="BUSM1"
               misc_feature
               1. 2540
               /note="assembly_fragment:01940

```

```

misc_feature      fragment_chain:1"
2641..85546
/note="assembly_fragment:01906
fragment_chain:1"
misc_feature      85647..100008
/note="assembly_fragment:02336
fragment_chain:1"
misc_feature      100109..106574
/note="assembly_fragment:06581
fragment_chain:1"
misc_feature      106675..126703
/note="assembly_fragment:03339.0"
126804..130027
/note="assembly_fragment:03339.1"
130128..133270
/note="assembly_fragment:04734"
133371..135771
/note="assembly_fragment:04895"
135872..136584
/note="assembly_fragment:06490"
138685..140983
/note="assembly_fragment:06532"
141084..147734
/note="assembly_fragment:06663"
BASE COUNT      45653 a 28122 c 28427 g 44527 t 1005 others
ORIGIN

```

```

Query Match      87.0%; Score 17.4; DB 1; Length 147734;
Best Local Similarity 94.7%; Pred. No. 74;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      2 TACTCTTACTTATTAAGT 20
Db 96725 TAGTCTTATTAAGT 96743

```

Search completed: April 6, 2003, 11:41:10
 Job time : 516.5 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 10:34:07 ; Search time 243 Seconds
(without alignments)
1069.066 Million cell updates/sec

Title: US-09-980-265-9
Perfect score: 20
Sequence: 1 ttatgtcttattatgatt 20

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 1994485 seqs, 6494577260 residues
Total number of hits satisfying chosen parameters: 3988970

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl_NoHTG:*
1: gb_ba:*
2: gb_in:*
3: gb_ov:*
4: gb_ov:*
5: gb_ov:*
6: gb_ph:*
7: gb_ph:*
8: gb_ph:*
9: gb_ph:*
10: gb_ph:*
11: gb_ph:*
12: gb_ph:*
13: gb_ph:*
14: gb_ph:*
15: gb_ph:*
16: gb_ph:*
17: gb_ph:*
18: gb_ph:*
19: gb_ph:*
20: gb_ph:*
21: gb_ph:*
22: gb_ph:*
23: gb_ph:*
24: gb_ph:*
25: gb_ph:*
26: gb_ph:*
27: gb_ph:*
28: gb_ph:*
29: gb_ph:*
30: gb_ph:*
31: gb_ph:*
32: gb_ph:*
33: gb_ph:*
34: gb_ph:*
35: gb_ph:*
36: gb_ph:*
37: gb_ph:*
38: gb_ph:*
39: gb_ph:*
40: gb_ph:*

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	20	100.0	20	5	AX057375 Sequence
2	18.4	92.0	20	5	AX057367 Sequence
3	18.4	92.0	20	5	AX057371 Sequence
4	18.4	92.0	20	5	AX057376 Sequence
5	18.4	92.0	20	5	AX057377 Sequence
6	18.4	92.0	20	5	AX057378 Sequence
7	18.4	92.0	20	5	AX057379 Sequence
8	18.4	92.0	65	5	AX485115 Sequence
9	18.4	92.0	749	7	AF402509 Sequence
10	18.4	92.0	25102	7	AF128394 Sequence
11	18.4	92.0	35468	2	AF098500 Sequence
12	18.4	92.0	45205	2	U41014 Sequence
13	18.4	92.0	105138	8	AF124730 Sequence
14	18.4	92.0	153094	8	CNS05TCH Sequence
15	18.4	92.0	165920	8	CNS05TCH Sequence
16	18.4	92.0	170892	8	CNS05TCH Sequence
17	18.4	92.0	174919	2	AC009370 Sequence
18	18.4	92.0	175867	2	AC009376 Sequence
19	18.4	92.0	178628	2	AC009385 Sequence
20	18.4	92.0	189413	8	AF129075 Sequence
21	18.4	92.0	196296	7	ATCHRIV10 Sequence
22	18.4	92.0	292226	2	AE003520 Sequence
23	18.4	92.0	324484	2	AE003584 Sequence
24	18.4	92.0	340000	8	HS21C049 Sequence
25	18.4	92.0	349050	1	AP003586 Sequence
26	18.4	90.0	81332	7	AP004959 Sequence
27	17.4	87.0	519	7	AF464739 Sequence
28	17.4	87.0	569	7	AF464744 Sequence
29	17.4	87.0	570	7	AF503678 Sequence
30	17.4	87.0	576	7	AF464745 Sequence
31	17.4	87.0	589	7	AF464746 Sequence
32	17.4	87.0	600	7	AF464746 Sequence
33	17.4	87.0	600	7	AF464750 Sequence
34	17.4	87.0	601	7	AF464748 Sequence
35	17.4	87.0	602	7	AF464741 Sequence
36	17.4	87.0	602	7	AF464749 Sequence
37	17.4	87.0	603	7	AF464742 Sequence
38	17.4	87.0	603	7	AF464743 Sequence
39	17.4	87.0	603	7	AF464751 Sequence
40	17.4	87.0	603	7	AF464752 Sequence
41	17.4	87.0	603	7	AF464753 Sequence
42	17.4	87.0	651	1	AB031217 Sequence
43	17.4	87.0	651	1	AB031221 Sequence
44	17.4	87.0	1955	8	AB052138 Sequence
45	17.4	87.0	7238	5	AX345350 Sequence

ALIGNMENTS

RESULT 1	AX057375	Sequence 9 from Patent WO0075304.	20 bp	DNA	linear	PAT 17-JAN-2001
LOCUS	AX057375					
DEFINITION	AX057375					
ACCESSION	AX057375.1	GI:12310116				
VERSION	AX057375.1					
KEYWORDS						
SOURCE						
ORGANISM						
REFERENCE						
AUTHORS						
TITLE						
JOURNAL						
FEATURES						

source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide"
BASE COUNT 4 a 1 c 3 g 12 t
ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTTATTAGTT 20
|||||
Db 1 TTAGTCTTAGTTATTAGTT 20

RESULT 2
AX057367 20 bp DNA linear PAT 17-JAN-2001
LOCUS AX057367
DEFINITION Sequence 1 from Patent WO0075304.
ACCESSION AX057367
VERSION AX057367.1 GI:12310108
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 1 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide"

BASE COUNT 5 a 0 c 3 g 12 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTTATTAGTT 20
|||||
Db 1 TTAGTCTTAGTTATTAGTT 20

RESULT 3
AX057371 20 bp DNA linear PAT 17-JAN-2001
LOCUS AX057371
DEFINITION Sequence 5 from Patent WO0075304.
ACCESSION AX057371
VERSION AX057371.1 GI:12310112
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 5 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide"

BASE COUNT 4 a 0 c 3 g 13 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTTATTAGTT 20
|||||
Db 1 TTAGTCTTAGTTATTAGTT 20

RESULT 4
AX057376 20 bp DNA linear PAT 17-JAN-2001
LOCUS AX057376
DEFINITION Sequence 10 from Patent WO0075304.
ACCESSION AX057376
VERSION AX057376.1 GI:12310117
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 10 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide"

BASE COUNT 3 a 1 c 3 g 13 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTTATTAGTT 20
|||||
Db 1 TTAGTCTTAGTTATTAGTT 20

RESULT 5
AX057377 20 bp DNA linear PAT 17-JAN-2001
LOCUS AX057377
DEFINITION Sequence 11 from Patent WO0075304.
ACCESSION AX057377
VERSION AX057377.1 GI:12310118
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 11 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide"

BASE COUNT 3 a 2 c 3 g 12 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTTATTAGTT 20
|||||
Db 1 TTAGTCTTAGTTATTAGTT 20

RESULT 6
AX057378

```
LOCUS       AX057378                20 bp    DNA                linear    PAT 17-JAN-2001
DEFINITION   Sequence 12 from Patent WO0075304.
ACCESSION    AX057378
VERSION      AX057378.1  GI:12310119
KEYWORDS
SOURCE       .
ORGANISM     synthetic construct.
              artificial construct.
REFERENCE    1 (bases 1 to 20)
AUTHORS      Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE        Immunostimulant oligonucleotide
JOURNAL      Patent: WO 0075304-A 12 14-DEC-2000;
              Aventis Pasteur (FR)
FEATURES
  source     1..20
              /organism="synthetic construct"
              /db_xref="taxon:32630"
              /note="oligonucleotide"
BASE COUNT   4 a          0 c          4 g          12 t
ORIGIN
Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY          1 TTAGTCTTAGTATTAGTT 20
            ||||| ||||| ||||| |||||
Db          1 TTAGTCTTAGTATTAGTT 20

RESULT 7
LOCUS       AX057379                20 bp    DNA                linear    PAT 17-JAN-2001
DEFINITION   Sequence 13 from Patent WO0075304.
ACCESSION    AX057379
VERSION      AX057379.1  GI:12310120
KEYWORDS
SOURCE       .
ORGANISM     synthetic construct.
              artificial construct.
REFERENCE    1 (bases 1 to 20)
AUTHORS      Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE        Immunostimulant oligonucleotide
JOURNAL      Patent: WO 0075304-A 13 14-DEC-2000;
              Aventis Pasteur (FR)
FEATURES
  source     1..20
              /organism="synthetic construct"
              /db_xref="taxon:32630"
              /note="oligonucleotide"
BASE COUNT   3 a          1 c          4 g          12 t
ORIGIN
Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY          1 TTAGTCTTAGTATTAGTT 20
            ||||| ||||| ||||| |||||
Db          1 TTAGTCTTAGTATTAGTT 20

RESULT 8
LOCUS       AX485115/c              65 bp    DNA                linear    PAT 16-AUG-2002
DEFINITION   Sequence 2415 from Patent WO02053728.
ACCESSION    AX485115
VERSION      AX485115.1  GI:22319399
KEYWORDS
SOURCE       .
ORGANISM     Candida albicans.
              Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
              Saccharomycetales; mitosporic Saccharomycetales; Candida.

LOCUS       AX057378                20 bp    DNA                linear    PAT 17-JAN-2001
DEFINITION   Sequence 12 from Patent WO0075304.
ACCESSION    AX057378
VERSION      AX057378.1  GI:12310119
KEYWORDS
SOURCE       .
ORGANISM     synthetic construct.
              artificial construct.
REFERENCE    1 (bases 1 to 20)
AUTHORS      Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE        Immunostimulant oligonucleotide
JOURNAL      Patent: WO 0075304-A 12 14-DEC-2000;
              Aventis Pasteur (FR)
FEATURES
  source     1..20
              /organism="synthetic construct"
              /db_xref="taxon:32630"
              /note="oligonucleotide"
BASE COUNT   4 a          0 c          4 g          12 t
ORIGIN
Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 6.1e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY          1 TTAGTCTTAGTATTAGTT 20
            ||||| ||||| ||||| |||||
Db          1 TTAGTCTTAGTATTAGTT 20

RESULT 9
LOCUS       AF402509/c              749 bp    DNA                linear    PLN 05-NOV-2001
DEFINITION   Primula cicutarillifolia ribosomal protein L16 (rpl16) gene, intron
              sequence; chloroplast gene for chloroplast product.
ACCESSION    AF402509
VERSION      AF402509.1  GI:16660214
KEYWORDS
SOURCE       .
ORGANISM     Primula cicutarillifolia.
              Chloroplast Primula cicutarillifolia
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              Asteridae; Ericales; Primulaceae; Primula.
              1 (bases 1 to 749)
              Mast,A.R., Kelso,S., Richards,A.J., Lang,D.J., Feller,D.M.S. and
              Conti,E.
              Phylogenetic Relationships in Primula L. and related Genera
              (Primulaceae) based on Noncoding Chloroplast DNA
              Int. J. Plant Sci. 162 (6), 1381-1400 (2001)
              Mast,A.R., Kelso,S., Richards,A.J., Lang,D.J., Feller,D.M.S. and
              Conti,E.
              Direct Submission
              Submitted (25-JUL-2001) Institute for Systematic Botany, University
              of Zurich, Zollikerstrasse 107, Zurich 8008, Switzerland
FEATURES
  source     1..749
              /organism="Primula cicutarillifolia"
              /organella="plastid:chloroplast"
              /db_xref="taxon:170920"
              <1..>749
              /gene="rpl16"
              /note="ribosomal protein L16"
              <1..>749
              /gene="rpl16"
BASE COUNT   305 a        110 c        129 g        205 t
ORIGIN
Query Match 92.0%; Score 18.4; DB 7; Length 749;
Best Local Similarity 95.0%; Pred. No. 2.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY          1 TTAGTCTTAGTATTAGTT 20
            ||||| ||||| ||||| |||||
Db          47 TTGGTCTTAGTATTAGTT 28

RESULT 10
LOCUS       T25H8                   25102 bp    DNA                linear    PLN 03-MAR-1999
DEFINITION   Arabidopsis thaliana BAC T25H8.
ACCESSION    AF128394
```

VERSION AF128394.1 GI:4325347
KEYWORDS
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana
Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots:
Rosidae: euroids II: Brassicales: Brassicaceae: Arabidopsis.
1 (bases 1 to 25102)
REFERENCE Washington University Genome Sequencing Center.
AUTHORS The A. thaliana Genome Sequencing Project
JOURNAL Unpublished (1997)
REFERENCE 2 (bases 1 to 25102)
AUTHORS Edwards, J. and Flagg, L.
TITLE The sequence of A. thaliana T25H8
JOURNAL Unpublished (1999)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (13-FEB-1999) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63108, USA
e-mail: r.watson@wustl.edu

MAPPING: Clones were assigned to the YAC map by hybridization by M. Lodhi, Cold Spring Harbor Laboratories, and fingerprinted by M. Marra, Washu, to pick the best candidates for sequencing.

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded or sequenced with an alternate chemistry; an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone

NEIGHBORING COSMID INFORMATION:

The 5' clone is T7M24, 200 bp overlap; 3' clone is T24M8, 200 bp overlap. Actual start of this clone is unknown, the first base of overlap is at base position 37383 of T7M24; actual end is at 35788 of T24M8.

NOTES:

Coding sequences below are predicted from computer analysis, using the program GeneFinder (P. Green and L. Hillier, ms in preparation). Location/Qualifiers

1. 25102

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="IV"

/map="unknown"

/clone="T25H8"

/complement(7415..10295)

/gene="T25H8.4"

/complement(join(7415..7864,8203..9163,10111..10295))

/gene="T25H8.4"

/note="contains similarity to phosphate transporters"

/codon_start=1

/evidence-not_experimental

/protein_id="AADI7346.1"

/db_xref="GI:4325348"

gene
CDS

gene
CDS
/translation="MVKSPFEIEDMLPEPSSSEEDYSADQKAYFEKSNESYH
MSRMSQTPLEKMTSISVLDGSMTERQIRQACQARSSLQIWESSNRKAMQRE
VARCKRIYAEAEHDEAVFEAVKSEVEIIGDTPAPVGRNRPKRREPTPSKY
OYLMELKEGTRYPHKEMTQIEIGIDVEYLIGLNLATFMSCQCGEKEQCILAT
LKHFHVEDSEKEERGICLYTFKKGVEYVLDIHLDTIFGFPGEELRODDEL
LSMTITINGPKRYPASRSKTSIRSDPRLDHQICATLPKPTGPGGELICIDQ
ALVFLTRKTRKGRKAKGRADSLITVLLDILGSLVROYAATLRSLICQVLEPSE
OPEDETEDEANHEEDADQGDDEIAQHKHGLVLELGNHRSKTKMGKKRTIPKK
SKSMAAOIKAFQSKPRPSPPKVRFSGFTSSAARRADRIDQPRASSFEPREVISL
15572..16980
/gene="T25H8.3"

join(15572..15841,16072..16288,16437..16604,16730..16980)

/gene="T25H8.3"

/codon_start=1

/evidence-not_experimental

/protein_id="AADI7348.1"

/db_xref="GI:4325350"

/translation="MVDNERACEKEESSLSRTVGSQDSNEETNANDLSFANPD
EEDERDDEYIVKSLDEKGTIVIEKVEEPTVKDIPMAIIMNINSKRLRMVY
LFLVSLIVLTKTEKGSVDNFCIRASADLFSKTEPMKRTSEPMKPSHTLDH
FNALAIPLRQFLIEKRGAVTGCRRCKVFKFTETLKGFTLEQINNVLDTEVSIP
RVEDVDKHDVVVDSMMRIGGREIRREEVYEDVQTRKAPTEEVPVAVPGHLL
LYDLMEXLDSINDKLNFAVALADY"

17953..20530

/gene="T25H8.2"

join(17953..19957,20012..20107,20178..20530)

/gene="T25H8.2"

/note="similar to Antirrhinum majus (garden snapdragon)

TNP2 protein (GB:X57297)"

/codon_start=1

/evidence-not_experimental

/protein_id="AADI7349.1"

/db_xref="GI:4325351"

/translation="MAGNYNKGSGGFADMYKRFDEVTGNLSAYVAGYEEFTFA
NSQPIVOSRGKFCPCSGYCKNEKHIIGRSYSLFSEFPEPDYVYWKSEELMD
ICTSYDRYSENEHEGVNVEDPYVMVNDAPFNNGYDNDVNDHNDYHDSGYON
VEEPRNHSNKTYDILEGANNPLDYGCEGOSLSASRLMINKAEYMKSELYVSYC
EMFTAFLEBGNQATTSHTQTEKLMRNILGIPHTIDVYKNCMLFKEDKEEDCFGC
AORMKPKDORRTKVPYSRWYLPADRLKRYOSHKRAAARWIAEHQSKGKGNHP
SDAAEMRYFOELHPFAEPINNVYGLTYDGNPNPMSRNSLAMPVILTPYNLPGMC
NMTIELFLTLINSNGNHPRASLDVFLQPLIEELKELMCTGVADYVLSQNFNLKAVL
LMTISDPAYSMISGNTTHGKISCPCVCMSESTFYLPGRKTCDFCDHRRFLPHGPS
RNKKDPLKGRDASSEYPRESTIGQVYTERLASYNPKTYDYGNGHKKRGYGE
HNMKESILMELSTWKDLNRHNDVHTERKFNELNNMTLGVGKSKDNLMDLI
EKYCSRPGILIDSTGKAPPPYKYLEAKQSLQCVKHDPVPPDMKSHDCHVEMRL
LPFIFAELDRNVHLALSGIRAFERDLSRTLOTSRVOILKONIVLIICNLEKIPPS
PEDVMEHLPIHLPEYAEELGGPVQYRWMPFEREFKLGKAKNKRYAAGSIYESYIND
EIAVSEHYFADHIOTKSR"

22613..24984

/gene="T25H8.1"

join(22613..23078,24243..24499,24634..24984)

/gene="T25H8.1"

/note="contains similarity to Petunia PTTA' (GB:AF009516)"

/codon_start=1

/evidence-not_experimental

/protein_id="AADI7347.1"

/db_xref="GI:4325349"

/translation="MPNAGRGSGRRKRTTPNVYIQRAGSTPAGRPSSLPQOYFTPA
ATVOVASIPGAGAGASSAPRYRNPPOOLRHSNORPOVDPLPPOETAOOP
PLSPQEPATSHSPSSCGNNROEGIPALVELQEDSVVANDILSVPRBAMCWVOY
YFNETVYKRLKDMYSTARTTRQOPWICETLMTGTCATVETBAOKRSRTYSKALSD
RNGICPHVHYSQPKSFOEIDOLEKLEKRLPHLGVFLIETHKSGSFDVDSKETAQ
AYQOVNRDLTALAEASAVDSGSRPELTLDTDTAIFLEKRPSSIIISLWPNIL
SLMSPFISCSPOKMIQCALIMD"

7623 a 4781 c 5019 g 7679 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 7; Length 25102;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 TTAGTCTTAGTTAGTTAGTT 20

Query

Db 23783 TTAGTCTTACTGACTTACT 23802

|||||

RESULT 11 35468 bp DNA linear INV 22-MAY-2002
AF098500
LOCUS
DEFINITION Caenorhabditis elegans cosmid F43B10, complete sequence.
ACCESSION AF098500
VERSION AF098500.1 GI:3786411
KEYWORDS HTG.
SOURCE Caenorhabditis elegans.
ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 35468)

AUTHORS Waterston, R.
TITLE Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium
JOURNAL Science 282 (5396), 2012-2018 (1998)
MEDLINE 99069613
PUBMED 9851916

REFERENCE 2 (bases 1 to 35468)

AUTHORS Fulton, R., Hawkins, J. and Rohlfing, T.
TITLE The sequence of C. elegans cosmid F43B10
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 35468)

AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (13-OCT-1998) Department of Genetics, Washington University, 444 Forest Park Avenue, St. Louis, Missouri 63108, USA

REFERENCE 4 (bases 1 to 35468)

AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA
REFERENCE 5 (bases 1 to 35468)

AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA

COMMENT Submitted by:

Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1RQ, England
email: tw@nematode.wustl.edu and jee@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one m13 subclone.

For a graphical representation of this cosmid sequence and its analysis see:
<http://www.wormbase.org/db/seq/sequence?name=F43B10;class=Sequence>

NEIGHBORING COSMID INFORMATION

The 5' cosmid is C06G1, 14200 bp overlap; the 3' cosmid is F22H10, 200 bp overlap. Actual start of this cosmid is at base position 197 of F43B10, actual end is at 3375 of F22H10.

NOTES:

Coding sequences below are the result of integration and manual review of the following data: computer analysis using the program GeneFinder (P. Green and U. Hillier, personal communication), the large scale EST projects of Yoji Kohara (http://www.ddb.jni.ac.jp/c-elegans/html/CE_INDEX.html) and The C. elegans ORFome cloning project (<http://wormfdb.dcel.harvard.edu/>), similarity to other proteins from Blastx analyses (<http://blast.wustl.edu/>), sequence conservation with C. briggsae using Jim Kent's WABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans Genbank submissions, and personal communications with C. elegans researchers. tRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

FEATURES

SOURCE

1. 35468
/organism="Caenorhabditis elegans"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="X"
/clone="F43B10"
complement(20574..22725)
/gene="F43B10.1"
/note="for a graphical representation of this gene see: <http://www.wormbase.org/db/seq/sequence?name=F43B10.1;class=Sequence>"
complement(join(20574..20880,21108..21251,21323..21503,21868..21942,22322..22367,22463..22725))
/gene="F43B10.1"
/note="coded for by the following C. elegans cDNAs: yk350f2.5"
/codon_start=1
/product="Hypothetical protein F43B10.1"
/protein_id="AAC67400.2"
/db_xref="GI:14389812"
/translation="MALLHLAPRLGDIFFLCVSIFSLVFLSLYLSHGHRARLIIFISVFRMLSLIMPPPPPMKVVVGRPLTAYELNNAIFFWHITMARGRKORLKVTRISIRFCHDKIVKMLTSPDRSSSGAEVLSLEDDDKPKSYKTPFMKLDLSHGSPDLRSDLPALITTPPKKNVNDLRSPETITTSADRNLVRLRRKSITESPRTNLGAGIESLPAYIAKKDNVNSQTLISWITFTSPHRRERKMGGGGDDAIRKMSQEVKKRRRIEBSGGGGRGVNAWIMAKRT"
complement(24967..34263)
/gene="F43B10.2"
/note="for a graphical representation of this gene see: <http://www.wormbase.org/db/seq/sequence?name=F43B10.2;class=Sequence>"
complement(join(24967..25126,25277..25463,25509..25533,25750..25963,26181..26287,26340..26520,26565..26740,26795..26853,26901..27089,27340..27450,27497..27727,27887..27950,29903..30015,30752..31138,31218..31327,32093..32175,32472..32631,32674..32806,33844..33953,33996..34135,34182..34263))
/gene="F43B10.2"
/note="contains similarity to human ALR (GB:AF003386); coded for by the following C. elegans cDNAs: yk12e10.3, yk12e10.5, yk350b9.3, yk350b9.5"
/codon_start=1
/product="Hypothetical protein F43B10.2"
/protein_id="AAC67399.1"
/db_xref="GI:3786412"
/translation="MISPSKFEVGEALRQKDVRFVSAKDNEERENVYLLGPTS SOKSELIDFLCNFYGVSDMDQRPHISNEFSETPRKPVQCYFNTKMSVRVILDMACGDYDGEFSLIHKMLLNKMKRLVIAVFSGLRMSHHEDELOOVSNTLPEHVRNSIVFRTASDGRPEPLNRRGLSECPFTINTSCTRKQWEDLNDHR RYKMSVNOEDSEMEKQVETTPITISGLPEYDDGICVQESGONSRSRSYSSSKSTV ILEVQSSRQOQVGRPTESPLPPLPLSNPPAKTPPPPKPIIRPPPIIPASPPR PPASEPVGIPLYSERQOQOQLRPTTAVSPKPSAPVPAALITPTSTTNALTN AMFQLOKOSVRYDVPYKSDRSRGTQTCQSLPPACVGRHRSVPDVHAYDETT PPVYLHVADRVSTVAYVWNEHPHSTNISTVSHSPRSQVGRDARLSRDRLRVYDAQ SSPADSOSEELRRMYSGRSGSVKPTGVTAARYVDTRNRYTAVGEOARRESYTSFT NDKRMRSRSEPSSEYTIINDPYISGRKYVSGCYETFRRSRTSLSHSGVPELISHS ENKYDQQLTDQNMERKSRILDYERRSRROSESRQYITDRIQISPERQDQYSGTY

gene

CDS

gene

gene

gene

CDS

gene

gene

gene

gene

gene

[illegible]

Query Match	92.0%;	Score 18.4;	DB 2;	Length 45205;
Best Local Similarity	95.0%;	Pred. No. 80;		
Matches	19;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;

Qy	1	TTAGTTCCTAGTTATTAGTT	20
Db	38063	TTAGTTCCTAGTTATTACTT	38082

RESULT	13
AF124730/c	
LOCUS	105138 bp DNA linear PRI 13-APR-2001
DEFINITION	Homo sapiens chromosome 21 clone PAC N2184 map 21q22.1, complete sequence.
ACCESSION	AF124730
VERSION	AF124730.2 GI:13621228
KEYWORDS	HTG.
SOURCE	Homo sapiens.
ORGANISM	Homo sapiens

REFERENCE
AUTHORS

Peik, K., S. Toyoda, A., Ishii, A., Yokoyama, C., Ochi, T., R., Soeta, E., Ohki, M., Takagi, T., Sakaki, Y., Tadain, S., Blechschmidt, K., Polley, A., Menzel, U., Delabar, J., Kumpf, K., Lehmann, R., Patterson, D., Reichwald, K., Rump, A., Schillhabel, M. B., Schudy, A., Zimmermann, W., Rosenthal, A., Kidoh, J., Kawasaki, K., Asakawa, S., Shintani, A., Sasaki, T., Nagamine, K., Mitsuyma, S., Antonarakis, S. E., Minoshima, S., Shimizu, N., Nordstiek, G., Hornischer, K., Brandt, P., Schafte, H., Schoen, O., Desario, A., Reichelt, J., Kauer, G., Bloeker, H., Ramsier, J., Beck, A., Klages, S., Henz, S., Riesselmann, L., Dandag, E., Haaf, T., Wehrmeyer, S., Borny, K., Gardiner, K., Nitzel, D., Francis, F., Lehrach, H., Reinhardt, R. and Yaspo, M. Laure.

TITLE
The DNA sequence of human chromosome 21
JOURNAL
Nature 405 (6784), 311-319 (2000)

REFERENCE

TITLE	Author
Direct Submission	Drescher, B., Weber, J., Schattevoy, R., Yaspo, M.-L. and Rosenthal, A.

REFERENCE

TITLE	AUTHORS
Direct Submission	Laugier, S., Daguau, E., Detard, V., Ull, R., Nourst, S., Drescher, B., Weber, J., Schatevov, R., Yaspo, M.-L. and Rosenthal, A

JOURNAL Submitted (13-APR-2001) Genome Analysis, Institute of Molecular Biotechnology, Beutenbergstr. 11, Jena 07745, Germany

COMMENT On Apr 13, 2001 this sequence version replaced gi:4262557.

FEATURES location/qualifiers

source

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q22.1"
/clone="PAC.N2184"
BASE COUNT      30268 a 20684 c 20685 g 33501 t
ORIGIN

```

Query Match	92.0%	Score 18.4	DB 8	Length 105138
Best Local Similarity	95.0%	Pred. No. 64		
Matches 19; Conservative	0	Mismatches 1	Indels 0	Gaps 0

```

QY      1  TTAGTCTTAGTTATTAGTT  20
          |||||
Db 20497 TTAGTCTTAGTTAATAGTT  20478

```

```
RESULT 14
CNS05TCH      153094 bp      DNA      linear      PRI 08-JUN-2001
LOCUS          Human chromosome 14 DNA sequence BAC R-241E13 of library RPCI-11
DEFINITION    from chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION     AL355098
VERSION       AL355098.3      GI:13677194
KEYWORDS      HTG; HTGS; ACTIVEFIN.
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE     1 (bases 1 to 153094)
               Hellig, R., Pett, J.L., Vico, V., Dasilva, C., Robert, C., Mincker, P.,
               Brothier, P., Catolico, L., Barbe, V., Pelletier, E., Artiguenave, F.,
               Levy, M., Eckenberg, R., Bruls, T., deBerardinis, V., Cruaud, C.,
               Gyapay, G., Saurin, W. and Weissbach, J.
               Sequencing of the human chromosome 14
               Unpublished
               2 (bases 1 to 153094)
               Genoscope.
               Direct Submission
               Submitted (07-JUN-2001) Genoscope - Centre National de Sequencage :
               BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
               Web : www.genoscope.cns.fr)
               On Apr 19, 2001 this sequence version replaced gi:9213458.
               ----- Genome Center
               Center: Genoscope / Centre National de Sequencage
               Center code: GS
               Web site: http://www.genoscope.cns.fr/
               Contact: SeqRef@genoscope.cns.fr
               -----
               The following BAC sequence is oriented from the T7 to the SP6 end.
               Upstream BAC (overlapping the T7 end) : C-2506J14
               Downstream BAC (overlapping the SP6 end) : R-76E12 (AC-AL163872).
               ----- Summary Statistics
               Assembly program: Phrap; version 2.0
               Quality coverage: 8.79x in Q20 bases; sum-of-contigs
               -----
               Overall quality chart :
               Range      :      bases
               0 - 9      :
               10 - 19   :
               20 - 29   :
               30 - 39   :      34
               40 - 49   :      1110
               50 - 59   :      4764
               60 - 69   :      4337
               70 - 79   :      8503
               80 - 89   :      30075
               90 - 99   :      104271
               -----
               Percentage of bases with a quality value >= 40 : 99 %.
               -----
               Location/Qualifiers
               1. 153094
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="14"
               /clone="R-241E13"
               /clone_lib="RPCI-11"
               BASE COUNT      43932 a      31270 c      32456 g      45436 t
ORIGIN
Query Match      92.0%: Score 18.4; DB 8; Length 153094;
Best Local Similarity 95.0%: Pred. No. 58;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 TTAGTCTTAGTATTAGTT 20
        ||||| ||||| |||||
Db 31203 TTAGTCTTAATATTAGTT 31222
```

```
RESULT 15
AC010232/c     165920 bp      DNA      linear      PRI 26-FEB-2002
LOCUS          Homo sapiens chromosome 5 clone CTC-316C21, complete sequence.
DEFINITION    AC010232
ACCESSION     AC010232
VERSION       AC010232.10     GI:18921277
KEYWORDS      HTG.
SOURCE        Homo sapiens.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE     1 (bases 1 to 165920)
               Doe Joint Genome Institute and Stanford Human Genome Center.
               Direct Submission
               Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
               Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
               3 (bases 1 to 165920)
               DOE Joint Genome Institute and Stanford Human Genome Center.
               Direct Submission
               Submitted (26-FEB-2002) DOE Joint Genome Institute, 2800 Mitchell
               Drive, Walnut Creek, CA 94598, USA
               On Feb 26, 2002 this sequence version replaced gi:15290321.
               Draft Sequence Produced by DOE Joint Genome Institute
               www.jgi.doe.gov
               -----
               Finishing completed at Stanford Human Genome Center
               www.shgc.stanford.edu
               Quality: Phrap Quality >=40 99.9% of Sequence;
               Estimated Total Number of Errors is 0.1.
               -----
               Location/Qualifiers
               1. 165920
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
               /chromosome="5"
               /clone="CTC-316C21"
               BASE COUNT      45866 a      32687 c      34867 g      52500 t
ORIGIN
Query Match      92.0%: Score 18.4; DB 8; Length 165920;
Best Local Similarity 95.0%: Pred. No. 57;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 TTAGTCTTAGTATTAGTT 20
        ||||| ||||| |||||
Db 50747 TTAGTATTAGTATTAGTT 50728
```

Search completed: April 6, 2003, 11:53:21
Job time : 336 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:21:32 ; Search time 151 Seconds
(without alignments)
298.278 Million cell updates/sec

Title: US-09-980-265-9

Perfect score: 1 ttagtcttagtatttagt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	22	AAF28880
2	19	95.0	20	22	AAF28871
3	18.4	92.0	20	22	AAF28872
4	18.4	92.0	20	22	AAF28876
5	18.4	92.0	20	22	AAF28881
6	18.4	92.0	20	22	AAF28882
7	18.4	92.0	20	22	AAF28883
8	18.4	92.0	20	22	AAF28884
9	18.4	92.0	749	24	ABQ21728

C	10	18.4	92.0	749	24	ABQ21729	Oligonucleotide fo
	11	17.4	87.0	7238	24	ABL32448	Human immune syste
	12	17.4	87.0	9507	24	ABN80243	Human chemically m
	13	16.8	84.0		20	AAF28873	Immunostimulatory
	14	16.8	84.0		20	AAF28874	Immunostimulatory
	15	16.8	84.0		20	AAF28875	Immunostimulatory
	16	16.8	84.0		20	AAF28876	Immunostimulatory
	17	16.8	84.0		20	AAF28877	Immunostimulatory
	18	16.8	84.0		20	AAF28878	Immunostimulatory
	19	16.8	84.0		20	AAF28879	Immunostimulatory
	20	16.8	84.0		20	AAF28880	Immunostimulatory
	21	16.8	84.0		20	AAF28881	Immunostimulatory
	22	16.8	84.0		20	AAF28882	Immunostimulatory
	23	16.8	84.0		20	AAF28883	Immunostimulatory
	24	16.8	84.0		20	AAF28884	Immunostimulatory
C	25	16.8	84.0	902	24	ABQ22435	Oligonucleotide fo
	26	16.8	84.0	1069	24	AAAC6432	Arabidopsis thalia
	27	16.8	84.0	1071	21	AAC39933	Arabidopsis thalia
	28	16.8	84.0	1240	21	AAC45636	Arabidopsis thalia
	29	16.8	84.0	1242	21	AAC40226	Arabidopsis thalia
C	30	16.8	84.0	1818	22	AAH13793	Human cDNA sequenc
	31	16.8	84.0	4741	23	ABL27984	Drosophila melanog
	32	16.8	84.0	5453	24	ABL70450	Chemically treated
	33	16.8	84.0	5453	24	AA561413	Human gene regulat
	34	16.8	84.0	5453	24	ABK31481	Signal transductio
	35	16.8	84.0	6759	24	ABL32739	Human immune syste
	36	16.8	84.0	7309	23	ABL29678	Drosophila melanog
	37	16.8	84.0	7353	24	ABL32072	Human immune syste
	38	16.8	84.0	7353	24	AAD28362	Human chemically t
	39	16.8	84.0	10682	24	ABN80085	Human chemically m
	40	16.8	84.0	11029	22	AA546414	Tumour suppressor
	41	16.8	84.0	11670	24	AA543325	Chemically treated
	42	16.8	84.0	13377	22	AA546476	Tumour suppressor
C	43	16.8	84.0	13377	24	ABL33463	Human immune syste
	44	16.8	84.0	13798	23	ABL8500	Drosophila melanog
	45	16.8	84.0	16258	24	ABL70376	Chemically treated

ALIGNMENTS

RESULT 1	
AAF28880	AAF28880 standard; DNA: 20 BP.
XX	
AC	AAF28880;
XX	
DT	09-MAY-2001 (first entry)
XX	
DE	Immunostimulatory oligonucleotide #9 as vaccine adjuvant.
XX	
KW	Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW	prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW	cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW	phosphorothioate; ss.
XX	
OS	Synthetic.
XX	
FH	Key
FT	modified_base
FT	1..20
FT	/*tag= a
FT	/note= "contain phosphorothioate internucleotide
FT	linkages"
XX	
FN	WO200075304-A1.
XX	
PD	14-DEC-2000.
XX	
PF	08-JUN-2000; 2000WO-FR01566.
XX	
PR	08-JUN-1999; 99FR-0007457.
PR	06-AUG-1999; 99FR-0010378.
XX	

PA (AVET) AVENTIS PASTEUR.
 XX
 PI Bachy M, Sodoyer R, Trannoy E;
 XX
 DR WPI: 2001-041317/05.
 XX
 PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
 PT vaccines for human use, induce lymphocyte proliferation and cytokine
 PT secretion -
 XX
 PS Example 4; Page 13; 30pp; French.
 XX
 CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
 CC immunostimulatory oligonucleotide of the invention which contains at
 CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
 CC oligonucleotides do not contain any CG dinucleotides in which C are
 CC unmethylated. The oligonucleotides are used as human immunostimulants
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
 CC They induce proliferation of human lymphocytes, induce secretion of
 CC cytokines, especially Interleukin-10 or Interferon-gamma and increase
 CC expression of the CD86 activation marker or the CD25 cytokine receptor
 CC on human B lymphocytes.
 CC
 SQ Sequence 20 BP; 4 A; 1 C; 3 G; 12 T; 0 other;
 Query Match 100.0%; Score 20; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 TTAGTCTTACTTAACTT 20
 DB 1 TTAGTCTTACTTAACTT 20
 RESULT 2
 AAF28871
 ID AAF28871 standard; DNA; 20 BP.
 XX
 AC AAF28871;
 XX
 DT 09-MAY-2001 (first entry)
 XX
 DE Generic immunostimulatory oligonucleotide as vaccine adjuvant.
 XX
 KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
 KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
 KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor; ss.
 XX
 OS Synthetic.
 OS
 PN WO200075304-A1.
 XX
 PD 14-DEC-2000.
 XX
 PE 08-JUN-2000; 2000WO-FR01566.
 XX
 PR 08-JUN-1999; 99FR-0007457.
 PR 06-AUG-1999; 99FR-0010378.
 XX
 PA (AVET) AVENTIS PASTEUR.
 XX
 PI Bachy M, Sodoyer R, Trannoy E;
 XX
 DR WPI: 2001-041317/05.
 XX
 PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
 PT vaccines for human use, induce lymphocyte proliferation and cytokine
 PT secretion -
 XX
 PS Claim 8; Page 17; 30pp; French.
 XX
 CC This sequence represents a generic example of an immunostimulatory
 CC oligonucleotide of the invention which contains at least one sequence

CC 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The oligonucleotides
 CC do not contain any CG dinucleotides in which C are unmethylated.
 CC The oligonucleotides are used as human immunostimulants and as adjuvants
 CC in therapeutic and prophylactic vaccines for human use. They induce
 CC proliferation of human lymphocytes, induce secretion of cytokines,
 CC especially interleukin-10 or interferon-gamma and increase expression
 CC of the CD86 activation marker or the CD25 cytokine receptor on human B
 CC lymphocytes.
 CC
 SQ Sequence 20 BP; 3 A; 1 C; 3 G; 12 T; 1 other;
 Query Match 95.0%; Score 19; DB 22; Length 20;
 Best Local Similarity 95.0%; Pred. No. 33;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 TTAGTCTTACTTAACTT 20
 DB 1 TTAGTCTTACTTAACTT 20
 RESULT 3
 AAF28872
 ID AAF28872 standard; DNA; 20 BP.
 XX
 AC AAF28872;
 XX
 DT 09-MAY-2001 (first entry)
 XX
 DE Immunostimulatory oligonucleotide #1 as vaccine adjuvant.
 XX
 KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
 KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
 KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
 KW phosphorothioate; ss.
 XX
 OS Synthetic.
 OS
 PN WO200075304-A1.
 XX
 PD 14-DEC-2000.
 XX
 PE 08-JUN-2000; 2000WO-FR01566.
 XX
 PR 08-JUN-1999; 99FR-0007457.
 PR 06-AUG-1999; 99FR-0010378.
 XX
 PA (AVET) AVENTIS PASTEUR.
 XX
 PI Bachy M, Sodoyer R, Trannoy E;
 XX
 DR WPI: 2001-041317/05.
 XX
 PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
 PT vaccines for human use, induce lymphocyte proliferation and cytokine
 PT secretion -
 XX
 PS Example 4; Page 13; 30pp; French.
 XX
 CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
 CC immunostimulatory oligonucleotide of the invention which contains at
 CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
 CC oligonucleotides do not contain any CG dinucleotides in which C are
 CC unmethylated. The oligonucleotides are used as human immunostimulants
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
 CC They induce proliferation of human lymphocytes, induce secretion of
 CC cytokines, especially interleukin-10 or interferon-gamma and increase
 CC expression of the CD86 activation marker or the CD25 cytokine receptor

```
CC on human B lymphocytes.
XX
SQ Sequence 20 BP; 5 A; 0 C; 3 G; 12 T; 0 other;
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTATTAGTT 20
    ||||| ||||| ||||| |||||
Db 1 TTAGTTATTAGTTATTAGTT 20

RESULT 4
AAF28876
ID AAF28876 standard; DNA: 20 BP.
XX
AC AAF28876;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #5 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /*tag="a
FT /note="contain phosphorothioate internucleotide
FT linkages"
XX
XX WO200075304-A1.
XX
XX 14-DEC-2000.
XX
XX 08-JUN-2000; 2000WO-FR01566.
XX
XX 08-JUN-1999; 99FR-0007457.
XX
XX 06-AUG-1999; 99FR-0010378.
XX
XX (AVET ) AVENTIS PASTEUR.
XX
XX
XX Bachy M, Sodoyer R, Trianny E;
XX
XX WPI: 2001-041317/05.
XX
XX
XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
XX PT vaccines for human use, induce lymphocyte proliferation and cytokine
XX PT secretion -
XX
XX Example 4; Page 13; 30pp; French.
XX
XX Oligonucleotides AAF28872-AAF28887 represent specific examples of an
XX CC immunostimulatory oligonucleotide of the invention which contains at
XX CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
XX CC oligonucleotides do not contain any CG dinucleotides in which C are
XX CC unethylated. The oligonucleotides are used as human immunostimulants
XX CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
XX CC They induce proliferation of human lymphocytes, induce secretion of
XX CC cytokines, especially interleukin-10 or interferon-gamma and increase
XX CC expression of the CD86 activation marker or the CD25 cytokine receptor
XX CC on human B lymphocytes.
XX
XX Sequence 20 BP; 4 A; 0 C; 3 G; 13 T; 0 other;
SQ
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 TTAGTCTTAGTATTAGTT 20
    ||||| ||||| ||||| |||||
Db 1 TTAGTTATTAGTTATTAGTT 20

RESULT 5
AAF28881
ID AAF28881 standard; DNA: 20 BP.
XX
AC AAF28881;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #10 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /*tag="a
FT /note="contain phosphorothioate internucleotide
FT linkages"
XX
XX WO200075304-A1.
XX
XX 14-DEC-2000.
XX
XX 08-JUN-2000; 2000WO-FR01566.
XX
XX 08-JUN-1999; 99FR-0007457.
XX
XX 06-AUG-1999; 99FR-0010378.
XX
XX (AVET ) AVENTIS PASTEUR.
XX
XX
XX Bachy M, Sodoyer R, Trianny E;
XX
XX WPI: 2001-041317/05.
XX
XX
XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
XX PT vaccines for human use, induce lymphocyte proliferation and cytokine
XX PT secretion -
XX
XX Example 4; Page 13; 30pp; French.
XX
XX Oligonucleotides AAF28872-AAF28887 represent specific examples of an
XX CC immunostimulatory oligonucleotide of the invention which contains at
XX CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
XX CC oligonucleotides do not contain any CG dinucleotides in which C are
XX CC unethylated. The oligonucleotides are used as human immunostimulants
XX CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
XX CC They induce proliferation of human lymphocytes, induce secretion of
XX CC cytokines, especially interleukin-10 or interferon-gamma and increase
XX CC expression of the CD86 activation marker or the CD25 cytokine receptor
XX CC on human B lymphocytes.
XX
XX Sequence 20 BP; 3 A; 1 C; 3 G; 13 T; 0 other;
SQ
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTATTAGTT 20
    ||||| ||||| ||||| |||||
Db 1 TTAGTTATTAGTTATTAGTT 20

RESULT 6
```

```
AAF28882
ID AAF28882 standard; DNA: 20 BP.
XX
AC AAF28882;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #11 as vaccine adjuvant.
XX
KM Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KM prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KM cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KM phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /tag= a
FT /note= "contain phosphorothioate internucleotide linkages"
FT
XX
XX WO200075304-A1.
XX
PD 14-DEC-2000.
XX
PF 08-JUN-2000; 2000WO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVER ) AVENTIS PASTEUR.
XX
PI Bachy M, Sodoyer R, Tranoy E;
XX
DR WPI: 2001-041317/05.
XX
XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
XX
SQ Sequence 20 BP: 3 A; 2 C; 3 G; 12 T; 0 other:
XX
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTATTAGTT 20
DB 1 TTAGTCTTAGTTCTTAGTT 20
XX
RESULT 7
AAF28883
ID AAF28883 standard; DNA: 20 BP.
XX
AC AAF28883;
XX
DT 09-MAY-2001 (first entry)
XX
```

```
DE Immunostimulatory oligonucleotide #12 as vaccine adjuvant.
XX
KM Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KM prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KM cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KM phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /tag= a
FT /note= "contain phosphorothioate internucleotide linkages"
FT
XX
XX WO200075304-A1.
XX
PD 14-DEC-2000.
XX
PF 08-JUN-2000; 2000WO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVER ) AVENTIS PASTEUR.
XX
PI Bachy M, Sodoyer R, Tranoy E;
XX
DR WPI: 2001-041317/05.
XX
XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
XX
SQ Sequence 20 BP: 3 A; 1 C; 4 G; 12 T; 0 other:
XX
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTATTAGTT 20
DB 1 TTAGTCTTAGTTGTTAGTT 20
XX
RESULT 8
AAF28884
ID AAF28884 standard; DNA: 20 BP.
XX
AC AAF28884;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #13 as vaccine adjuvant.
XX
KM Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KM prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KM cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KM phosphorothioate; ss.
XX
```

```

OS Synthetic.
XX Key Location/Qualifiers
FH modified_base 1..20
FT /*tag= a
FT /note= "contain phosphorothioate internucleotide
FT linkages"
XX
XX WO200075304-A1.
XX
XX 14-DEC-2000.
XX
XX 08-JUN-2000; 2000WO-FR01566.
XX
XX 08-JUN-1999; 99FR-0007457.
XX 06-AUG-1999; 99FR-0010378.
XX
XX (AVET ) AVENTIS PASTEUR.
XX
XX Bachy M, Sodoier R, Tranoy E;
XX WPI; 2001-041317/05.
XX
XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
XX vaccines for human use, induce lymphocyte proliferation and cytokine
XX secretion -
XX
XX Example 4; Page 13; 30pp; French.
XX
XX Oligonucleotides AAF28672-AAF28867 represent specific examples of an
XX immunostimulatory oligonucleotide of the invention which contains at
XX least one sequence 5'-TTNIN2PT-3' where N1 and N2 are A, T, C or G. The
XX oligonucleotides do not contain any CG dinucleotides in which C are
XX unmethylated. The oligonucleotides are used as human immunostimulants
XX and as adjuvants in therapeutic and prophylactic vaccines for human use.
XX They induce proliferation of human lymphocytes, induce secretion of
XX cytokines, especially interleukin-10 or interferon-gamma and increase
XX expression of the CD86 activation marker or the CD25 cytokine receptor
XX on human B lymphocytes.
XX
XX Sequence 20 BP; 4 A; 0 C; 4 G; 12 T; 0 other;
XX
XX Query Match 92.0%; Score 18.4; DB 22; Length 20;
XX Best Local Similarity 95.0%; Pred. No. 59;
XX Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 TTAGTCTTAGTTATTAGT 20
XX ||||| ||||| |||||
XX Db 1 TTAGTCTTAGTTATTAGT 20
XX
XX RESULT 9
XX ABO21728
XX ID ABO21728 standard; DNA; 749 BP.
XX
XX AC ABO21728;
XX
XX 12-JUL-2002 (first entry)
XX
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 8319.
XX
XX Human; cytosine methylation; 5'-CPG-3'; uracil; cytosine; diagnosis;
XX drug; side effect; cancer; central nervous system; cardiovascular;
XX gastrointestinal; respiratory system; single nucleotide polymorphism;
XX SNP; cell differentiation; ds.
XX
XX Homo sapiens.
XX
XX OS
XX
XX WO200218632-A2.
XX
XX PD 07-MAR-2002.
XX
XX 01-SEP-2001; 2001WO-EPI0074.
XX
XX PF

```

```

XX
XX 01-SEP-2000; 2000DE-1043826.
XX 05-SEP-2000; 2000DE-1044543.
XX
XX PA (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K, Guetig D;
XX WPI; 2002-371829/40.
XX
XX Determining the degree of cytosine methylation in genomic DNA, useful
XX for diagnosis and prognosis, comprises selective hybridization of
XX amplicons from chemically treated DNA -
XX
XX Claim 12; 56pp + Sequence Listing; 56pp; German.
XX
XX This invention describes a novel method for determining the degree of
XX methylation of a particular cytosine in a motif 5'-CPG-3', present in a
XX genomic sample of DNA. The sample is treated chemically to convert
XX cytosine (C) but not methylated C, to uracil, then part of the genomic
XX DNA that contains the target C is amplified to form a labeled amplicon.
XX The amplicon is hybridised to two classes, each with at least one
XX member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
XX and the degree of hybridisation to both classes is determined from the
XX label on the amplicon. From the ratio of labels hybridised to the two
XX classes of oligomers, the degree of methylation is calculated. The method
XX is used: (i) for diagnosis and/or prognosis of side effects of
XX therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
XX of the central nervous, cardiovascular, gastrointestinal and respiratory
XX systems etc., particularly by detecting mutations or single nucleotide
XX polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
XX types and for investigating cell differentiation. The method allows the
XX methylation status of many C residues to be determined simultaneously.
XX ABO13410-ABO54121 represent genomic DNA sequences used to illustrate the
XX method for determining the degree of cytosine methylation described in
XX the disclosure of the invention.
XX
XX Sequence 749 BP; 173 A; 54 C; 199 G; 322 T; 1 other;
XX
XX Query Match 92.0%; Score 18.4; DB 24; Length 749;
XX Best Local Similarity 95.0%; Pred. No. 52;
XX Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX QY 1 TTAGTCTTAGTTATTAGT 20
XX ||||| ||||| |||||
XX Db 347 TTAGTCTTAGTTATTAGT 366
XX
XX RESULT 10
XX ABO21729/c
XX ID ABO21729 standard; DNA; 749 BP.
XX
XX AC ABO21729;
XX
XX 12-JUL-2002 (first entry)
XX
XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 8320.
XX
XX Human; cytosine methylation; 5'-CPG-3'; uracil; cytosine; diagnosis;
XX drug; side effect; cancer; central nervous system; cardiovascular;
XX gastrointestinal; respiratory system; single nucleotide polymorphism;
XX SNP; cell differentiation; ds.
XX
XX Homo sapiens.
XX
XX OS
XX
XX WO200218632-A2.
XX
XX PD 07-MAR-2002.
XX
XX 01-SEP-2001; 2001WO-EPI0074.
XX
XX 01-SEP-2000; 2000DE-1043826.
XX 05-SEP-2000; 2000DE-1044543.
XX
XX PR

```

```
XX (EPiG-) EPIGENOMICS AG.
PA
XX
PI Olek A, Piepenbrock C, Berlin K, Guetig D;
XX
XX WPI: 2002-371829/40.
DR
XX
PT Determining the degree of cytosine methylation in genomic DNA, useful
PT for diagnosis and prognosis, comprises selective hybridization of
PT amplicons from chemically treated DNA -
XX
PS Claim 12: 56pp + Sequence Listing: 56pp; German.
XX
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one
CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
CC and the degree of hybridisation to both classes is determined from the
CC label on the amplicon. From the ratio of labels hybridised to the two
CC classes of oligomers, the degree of methylation is calculated. The method
CC is used: (i) for diagnosis and/or prognosis of side effects of
CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
CC of the central nervous, cardiovascular, gastrointestinal and respiratory
CC systems etc., particularly by detecting mutations or single nucleotide
CC polymorphisms (SNP's) and (ii) for differentiation of cell or tissue
CC types and for investigating cell differentiation. The method allows the
CC methylation status of many C residues to be determined simultaneously.
CC ABQJ3410-ABQ54121 represent genomic DNA sequences used to illustrate the
CC method for determining the degree of cytosine methylation described in
CC the disclosure of the invention.
XX
SQ Sequence 749 BP; 322 A; 199 C; 54 G; 173 T; 1 other;
XX
Query Match 92.0%; Score 18.4; DB 24; Length 749;
Best Local Similarity 95.0%; Pred. No. 52;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTTCCTTAGTTATTAGTT 20
DB 403 TTAGTTTCTTAGTTATTAGTT 384
RESULT 11
ABL32448
ID ABL32448 standard; DNA; 7238 BP.
XX
AC ABL32448;
XX
XX 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 421.
XX
XX Human; immune system disease; cytosine methylation; antiasthmatic;
XX antiarteriosclerotic; antihaemic; cytosstatic; noctropic;
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX antirheumatic; antiarthritic; antidiabetic; arteriosclerosis; anaemia;
XX antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
XX gene; ds.
XX
XX Homo sapiens.
XX
XX WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP07537.
XX
XX 30-JUN-2000; 2000DE-1032529.
XX
PR
```

```
PR 01-SEP-2000; 2000DE-1043826.
XX
XX (EPiG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI: 2002-130909/17.
DR
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
XX
PS Claim 1; SEQ ID NO 421; 32pp + Sequence Listing: German.
XX
XX The present invention provides a number of human immune system associated
XX genes which are modified by the methylation of cytosines. The sequences
XX can be used in the diagnosis and treatment of immune system disorders,
XX including eye diseases such as retinopathy, neovascular glaucoma and
XX macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
XX leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
XX rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
XX diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 7238 BP; 2414 A; 84 C; 1401 G; 3339 T; 0 other;
XX
Query Match 87.0%; Score 17.4; DB 24; Length 7238;
Best Local Similarity 94.7%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTTCCTTAGTTATTAGT 19
DB 6286 TTAGTTTCTTAGTTATTAGT 6304
RESULT 12
ABN80243
ID ABN80243 standard; DNA; 9507 BP.
XX
XX ABN80243;
XX
XX 15-JUL-2002 (first entry)
XX
DE Human chemically modified disease associated gene SEQ ID NO 260.
XX
XX Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
XX heart disease; epilepsy; epistone deacetylation; muscular dystrophy;
XX dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
XX antidiabetic; cytosstatic; anticonvulsant; ds.
XX
XX
XX Homo sapiens.
XX
XX Synthetic.
XX
XX WO200200927-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-EP07536.
XX
XX 30-JUN-2000; 2000DE-1032529.
XX
XX 01-SEP-2000; 2000DE-1043826.
XX
XX (EPiG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI: 2002-130908/17.
XX
XX Novel nucleic acid useful for diagnosis and therapy of diseases
XX associated with development genes such as diabetes, comprises a
XX sequence of a segment of chemically pretreated DNA of genes associated
XX with development -
XX
XX Claim 1; SEQ ID NO 260; 27pp; English.
XX
PS
```


XX The invention relates to a nucleic acid (I) comprising a sequence at
CC least 18 bases in length of a segment of chemically pretreated DNA (II)
CC of genes associated with development selected from 87 genes listed in
CC the specification such as ACCPN, ADFN, or AFD1 and comprising one of 350
CC sequences (ABN79984-ABN80333) or their complements. The invention is
CC useful for the diagnosis or therapy of diseases associated with
CC developmental genes, in particular disease related to homeobox containing
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
CC associated with congenital heart disease, epilepsy, diseases related to
CC histone deacetylation, Currarino syndrome, diseases related with the
CC development of the brain and limb girdle muscular dystrophy and dwarfism.
CC Oligomers specific to each of the genes are useful for detecting the
CC methylation state of all CpG dinucleotides within the 350 sequences or
CC (II) and their complementary sequences, as primer oligonucleotides for
CC the amplification of the 350 sequences, (II) and/or their complements and
CC as oligomer probes for detecting the cytosine methylation state and/or
CC single nucleotide polymorphisms (SNPs).
CC Note: The sequence data for this patent did not form part of the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.

SO Sequence 9507 BP; 2378 A; 326 C; 2397 G; 4406 T; 0 other;

Query Match 87.0%; Score 17.4; DB 24; Length 9507;
Best Local Similarity 94.7%; Pred. No. 1.3e+02;
Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTATTACT 19
||||| |||||||
DB 106 TTAGTTTATTAGTTTACT 124

RESULT 13

AA28873
ID AAF28873 standard; DNA; 20 BP.

AC AAF28873;

DT 09-MAY-2001 (first entry)

XX Immunostimulatory oligonucleotide #2 as vaccine adjuvant.

KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;

KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;

KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;

OS Synthetic.

XX Key Location/Qualifiers

FT modified_base 1..20

FT /tag= a

FT /note= "contain phosphorothioate internucleotide

FT linkages"

XX WO200075304-A1.

XX 14-DEC-2000.

XX 08-JUN-2000; 2000WO-FR01566.

XX 08-JUN-1999; 99FR-0007457.

XX 06-AUG-1999; 99FR-0010378.

XX (AVET) AVENTIS PASTEUR.

XX Bachy M, Sodoyer R, Tranoy E;

XX WPI: 2001-041317/05.

XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in

PT vaccines for human use, induce lymphocyte proliferation and cytokine

PT secretion -
XX Example 4; Page 13; 30pp; French.

CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNN2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unethyalted. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.

SO Sequence 20 BP; 4 A; 0 C; 3 G; 13 T; 0 other;

Query Match 84.0%; Score 16.8; DB 22; Length 20;
Best Local Similarity 90.0%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTATTACT 20
||||| |||||||
DB 1 TTAGTTATTAGTTTACT 20

RESULT 14

AA28874
ID AAF28874 standard; DNA; 20 BP.

AC AAF28874;

DT 09-MAY-2001 (first entry)

XX Immunostimulatory oligonucleotide #3 as vaccine adjuvant.

KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;

KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;

KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;

OS Synthetic.

XX Key Location/Qualifiers

FT modified_base 1..20

FT /tag= a

FT /note= "contain phosphorothioate internucleotide

FT linkages"

XX WO200075304-A1.

XX 14-DEC-2000.

XX 08-JUN-2000; 2000WO-FR01566.

XX 08-JUN-1999; 99FR-0007457.

XX 06-AUG-1999; 99FR-0010378.

XX (AVET) AVENTIS PASTEUR.

XX Bachy M, Sodoyer R, Tranoy E;

XX WPI: 2001-041317/05.

XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in

PT vaccines for human use, induce lymphocyte proliferation and cytokine

PT secretion -

XX Example 4; Page 13; 30pp; French.

CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an

CC immunostimulatory oligonucleotide of the invention which contains at

CC least one sequence 5'-TTNN2TT-3' where N1 and N2 are A, T, C or G. The

CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
XX
SQ Sequence 20 BP; 4 A; 1 C; 3 G; 12 T; 0 other;
Query Match 84.0%; Score 16.8; DB 22; Length 20;
Best Local Similarity 90.0%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TTAGTCTTAGTTATTAGTT 20
||||| ||||| |||||
Db 1 TTAGTTATTAGTTGTTAGTT 20
RESULT 15
AAF28875
ID AAF28875 standard; DNA: 20 BP.
XX
AC AAF28875;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #4 as vaccine adjuvant.
XX
KM Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KM prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KM cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KM phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /tag- a
FT /note- "contain phosphorothioate Internucleotide
linkages"
XX
PN WO200075304-A1.
XX
PD 14-DEC-2000.
XX
PF 08-JUN-2000; 2000WO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVET) AVENTIS PASTEUR.
XX
PI Bachy M, Sodey R, Tranoy E;
XX
DR WPI; 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTN1N2T-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.

XX
SQ Sequence 20 BP; 4 A; 0 C; 4 G; 12 T; 0 other;
Query Match 84.0%; Score 16.8; DB 22; Length 20;
Best Local Similarity 90.0%; Pred. No. 2.8e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TTAGTCTTAGTTATTAGTT 20
||||| ||||| |||||
Db 1 TTAGTTATTAGTTGTTAGTT 20
Search completed: April 6, 2003, 12:05:22
Job time : 152 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:12:37 ; Search time 45.25 Seconds
(without alignments)
135.548 Million cell updates/sec

Title: US-09-980-265-9

Perfect score: 20

Sequence: 1 ttactcttagttagtgc 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.8	84.0	59065	US-09-813-817-3	Sequence 3, Appli
2	16.8	84.0	59065	US-09-978-197-3	Sequence 3, Appli
3	15.8	79.0	29793	US-09-302-812-38	Sequence 38, Appli
4	15.8	79.0	29793	US-09-511-477-38	Sequence 38, Appli
5	15.8	79.0	29793	US-09-511-507-38	Sequence 38, Appli
6	15.2	76.0	403	US-08-565-166-30	Sequence 30, Appli
7	15.2	76.0	1203	US-09-122-443-3	Sequence 3, Appli
8	15.2	76.0	1578	US-09-134-001C-677	Sequence 677, App
9	15.2	76.0	1602	US-08-770-544-3	Sequence 3, Appli
10	15.2	76.0	1857	US-09-299-378-3	Sequence 3, Appli
11	15.2	76.0	2823	US-08-398-008A-1	Sequence 1, Appli
12	15.2	76.0	2823	US-08-893-333-1	Sequence 1, Appli
13	15.2	76.0	5049	US-08-336-345-2	Sequence 2, Appli
14	15.2	76.0	5049	US-08-647-655-2	Sequence 2, Appli
15	15.2	76.0	5935	US-09-178-973B-17	Sequence 17, Appli
16	15.2	76.0	5935	US-09-419-568F-29	Sequence 29, Appli
17	15.2	76.0	5935	US-09-354-243B-29	Sequence 29, Appli
18	14.8	74.0	20	US-08-588-821-76	Sequence 76, Appli
19	14.8	74.0	20	US-08-915-214-76	Sequence 76, Appli
20	14.8	74.0	20	US-09-005-532-76	Sequence 76, Appli
21	14.8	74.0	200	US-08-875-972-27	Sequence 27, Appli
22	14.8	74.0	292	US-08-644-664B-41	Sequence 41, Appli
23	14.8	74.0	292	US-08-761-277A-41	Sequence 41, Appli
24	14.8	74.0	348	US-08-220-606B-54	Sequence 54, Appli
25	14.8	74.0	437	US-09-397-787-315	Sequence 315, App
26	14.8	74.0	463	US-09-385-982-407	Sequence 407, App
27	14.8	74.0	626	US-09-328-111-519	Sequence 519, App

C	28	14.8	74.0	627	4	US-09-385-982-186	Sequence 186, App
	29	14.8	74.0	1753	6	5225348-2	Patent No. 5225348
	30	14.8	74.0	2128	2	US-08-371-377-16	Sequence 16, Appli
	31	14.8	74.0	3073	1	US-07-688-352C-31	Sequence 31, Appli
	32	14.8	74.0	3073	2	US-08-474-379C-31	Sequence 31, Appli
	33	14.8	74.0	3073	3	US-09-146-249A-31	Sequence 31, Appli
	34	14.8	74.0	3073	3	US-08-206-188B-31	Sequence 31, Appli
	35	14.8	74.0	3073	5	PCR-US91-02714-30	Sequence 30, Appli
	36	14.8	74.0	4695	4	US-09-309-572-9	Sequence 9, Appli
	37	14.8	74.0	4695	6	5225348-3	Patent No. 5225348
	38	14.8	74.0	111282	4	US-09-754-250-3	Sequence 3, Appli
	39	14.8	74.0	168575	4	US-09-426-290-1	Sequence 1, Appli
	40	14.8	74.0	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	41	14.4	72.0	5010	3	US-09-161-244-1	Sequence 442, App
	42	14.2	71.0	362	4	US-09-605-785-442	Sequence 442, App
	43	14.2	71.0	362	4	US-09-439-313-442	Sequence 442, App
	44	14.2	71.0	362	4	US-09-352-616A-442	Sequence 46, Appli
	45	14.2	71.0	427	4	US-09-556-877-46	

ALIGNMENTS

```
RESULT 1
US-09-813-817-3
; Sequence 3, Application US/09813817
; Patent No. 6340583
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001178
; CURRENT APPLICATION NUMBER: US/09/813,817
; CURRENT FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-813-817-3

Query Match      84.0% Score 16.8; DB 4; Length 59065;
Best Local Similarity 90.0%; Pred. No.33;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1 TTAGTCTTACTTATTAGTT 20
      11 ||||| |||||
DB 15314 TTGTCTTGTGTATTAGTT 15333

RESULT 2
US-09-978-197-3
; Sequence 3, Application US/09978197
; Patent No. 6403353
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: C1001178
; CURRENT APPLICATION NUMBER: US/09/978,197
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 09/813,817
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-978-197-3
```

Query Match 84.0%; Score 16.8; DB 4; Length 59065;
Best Local Similarity 90.0%; Pred. No. 32;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTCTTACTTACTT 20
Db 15314 TTGTCTTCTTACTTACTT 15333

RESULT 3

US-09-302-812-38
; Sequence 38, Application US/09302812B
; Patent No. 6335148
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AM, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOSYDROLASE (PARG) ENZ
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIVE
; FILE REFERENCE: NIAD 201
; CURRENT APPLICATION NUMBER: US/09/302,812B
; CURRENT FILING DATE: 1999-04-30
; EARLIER APPLICATION NUMBER: 60/083,768
; PRIOR FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 38
; LENGTH: 29793
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
US-09-302-812-38

Query Match 79.0%; Score 15.8; DB 4; Length 29793;
Best Local Similarity 89.5%; Pred. No. 86;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAGTCTTACTTACTT 20
Db 20450 TAGCTCTTACTTACTT 20468

RESULT 4

US-09-511-477-38
; Sequence 38, Application US/09511477
; Patent No. 6337202
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AME, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOSYDROLASE (PARG) EN
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREACTIV
; TITLE OF INVENTION: THEREWITH
; FILE REFERENCE: NIAD 201
; CURRENT APPLICATION NUMBER: US/09/511,477
; CURRENT FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/302,812
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 38
; LENGTH: 29793
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
US-09-511-477-38

Query Match 79.0%; Score 15.8; DB 4; Length 29793;
Best Local Similarity 89.5%; Pred. No. 86;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAGTCTTACTTACTT 20
Db 20450 TAGCTCTTACTTACTT 20468

RESULT 5

US-09-511-507-38
; Sequence 38, Application US/09511507
; Patent No. 6395543
; GENERAL INFORMATION:
; APPLICANT: JACOBSON, Myron K.
; APPLICANT: JACOBSON, Elaine L.
; APPLICANT: AME, Jean-Christophe
; APPLICANT: LIN, Winston
; TITLE OF INVENTION: GENES ENCODING SEVERAL POLY(ADP-RIBOSE) GLYCOSYDROLASE (PARG)
; TITLE OF INVENTION: THE PROTEINS AND FRAGMENTS THEREOF, AND ANTIBODIES IMMUNOREAC
; FILE REFERENCE: NIAD 201
; CURRENT APPLICATION NUMBER: US/09/511,507
; CURRENT FILING DATE: 2000-02-23
; PRIOR APPLICATION NUMBER: 09/302,812
; PRIOR FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 38
; LENGTH: 29793
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; FEATURE:
US-09-511-507-38

Query Match 79.0%; Score 15.8; DB 4; Length 29793;
Best Local Similarity 89.5%; Pred. No. 86;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAGTCTTACTTACTT 20
Db 20450 TAGCTCTTACTTACTT 20468

RESULT 6

US-08-569-166-30
; Sequence 30, Application US/08569166
; Patent No. 5830722
; GENERAL INFORMATION:
; APPLICANT: NICOLAS, LUC
; APPLICANT: CHARLES, JEAN-FRANCOIS
; APPLICANT: DELECLUSE, ARMELE
; APPLICANT: BARLOY, FREDERIQUE
; TITLE OF INVENTION: CLOSTRIDIUM BIFERMENTANS DNA FRAGMENT
; TITLE OF INVENTION: BEARING GENES CODING FOR PROTEINS LINKED TO AN
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/569,166
; FILING DATE: 05-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR94/00768
; FILING DATE: 24-JUN-1994
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: FR 93/07795
FILING DATE: 25-JUN-1993
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-106-0 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 30:
SEQUENCE CHARACTERISTICS:
LENGTH: 403 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-569-166-30

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 2; Length 403;
Pred. No. 1.8e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTATTAGTT 20
Db 99 TTACATCTTGTATTAGTT 118

RESULT 7
US-09-122-443-3/C
Sequence 3, Application US/09122443
Patent No. 6060284
GENERAL INFORMATION:
APPLICANT: Bazan, J. Fernando
TITLE OF INVENTION: MAMMALIAN CYTOKINE; RELATED REAGENTS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: DNAX Research Institute
STREET: 901 California Avenue
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/122,443
FILING DATE: 24-JUL-1998
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/053,765
FILING DATE: 25-JUL-1997
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34,090
REFERENCE/DOCKET NUMBER: DX0758K1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650)852-9196
TELEFAX: (650)496-1200
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1203 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 113..700
FEATURE:
NAME/KEY: mat_peptide

LOCATION: 176..700
US-09-122-443-3

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 3; Length 1203;
Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTATTAGTT 20
Db 1177 TTACTCTTAGTACTAGAT 1158

RESULT 8
US-09-134-001C-677
Sequence 677, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 677
LENGTH: 1578
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-677

Query Match
Best Local Similarity 76.0%; Score 15.2; DB 4; Length 1578;
Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTATTAGTT 20
Db 796 TTTGTATTAGTTTGT 815

RESULT 9
US-08-770-544-3
Sequence 3, Application US/08770544
Patent No. 5907085
GENERAL INFORMATION:
APPLICANT: Gonsalves, Dennis
TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND
TITLE OF INVENTION: THEIR USES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/770,544
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60009008
FILING DATE: 21-DEC-1995
ATTORNEY/AGENT INFORMATION:

```

: NAME: Goldman, Michael L.
: REGISTRATION NUMBER: 30,727
: REFERENCE/DOCKET NUMBER: 19603/621
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (716) 263-1304
: TELEFAX: (716) 263-1600
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1602 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: US-08-770-544-3

Query Match          76.0%; Score 15.2; DB 2; Length 1602;
Best Local Similarity 85.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0

QY      1 TTAGTTCCTAGTTATTAGTT 20
        ||| |||| |||| |||| |||
Db       1703 TTAATCTAAGTTATTAAAT 1722

RESULT 11
US-08-398-008A-1
: Sequence 1, Application US/08398008A
: Patent No. 5665588
: GENERAL INFORMATION:
: APPLICANT: Korblyuh, Jacki
: TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Gilbreth & Adler, P.C.
: STREET: 8011 Candle Lane
: CITY: Houston
: STATE: Texas
: COUNTRY: USA
: ZIP: 77071

```

```

1  COMPUTER READABLE FORM:
2  MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
3  COMPUTER: MACINTOSH IIcl
4  OPERATING SYSTEM: Macintosh
5  SOFTWARE: Microsoft Word 5.1
6  CURRENT APPLICATION DATA:
7  APPLICATION NUMBER: US/08/398,008A
8  FILING DATE: March 2, 1995
9  CLASSIFICATION: 435
10 PRIOR APPLICATION DATA:
11 APPLICATION NUMBER: 08/126,501
12 FILING DATE: 24-SEP-1993
13 ATTORNEY/AGENT INFORMATION:
14 NAME: Adler, Dr. Benjamin Aaron
15 REGISTRATION NUMBER: 35,423
16 REFERENCE/DOCKET NUMBER: D5705CIP
17 TELECOMMUNICATION INFORMATION:
18 TELEPHONE: (713) 777-2321
19 TELEFAX: (713) 777-6908
20 TELEX:
21 INFORMATION FOR SEQ ID NO: 1:
22 SEQUENCE CHARACTERISTICS:
23 LENGTH: 2823 base pairs
24 TYPE: nucleic acid
25 STRANDEDNESS: double-stranded
26 TOPOLOGY: unknown
27 MOLECULE TYPE: cDNA
28 HYPOTHEITICAL: no
29 ANTI-SENSE: no
30 US-08-398-008A-1
31
32 Query Match 76.0%; Score 15.2; DB 1; Length 2823;
33 Best Local Similarity 85.0%; Pred. No. 1.7e+02;
34 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
35
36 QY 1 TTAGTCTTACTTATTACTT 20
37 ||||| | | |||||
38 DB 2654 TTAGTTTAAATTATTACTT 2673
39
40 RESULT 12
41 US-08-893-333-1
42 Sequence 1, Application US/08893333A
43 Patent No. 5981705
44 GENERAL INFORMATION:
45 APPLICANT: Korbiduth, Jacki
46 TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated Protein
47 FILE REFERENCE: D5705CIP/D
48 CURRENT APPLICATION NUMBER: US/08/893,333A
49 CURRENT FILING DATE: 1997-07-16
50 NUMBER OF SEQ ID NOS: 17
51 SEQ ID NO 1
52 LENGTH: 2823
53 TYPE: DNA
54 ORGANISM: Homo sapiens
55 FEATURE:
56 LOCATION: 190..1953
57 OTHER INFORMATION: CDS
58 US-08-893-333-1
59
60 Query Match 76.0%; Score 15.2; DB 2; Length 2823;
61 Best Local Similarity 85.0%; Pred. No. 1.7e+02;
62 Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
63
64 OY 1 TTAGTCTTACTTATTACTT 20
65 ||||| | | |||||
66 DB 2654 TTAGTTTAAATTATTACTT 2673
67
68 RESULT 13
69 US-08-336-345-2
70 Sequence 2, Application US/08336345
71 Patent No. 5814510

```

```

: GENERAL INFORMATION:
: APPLICANT: Parrish, Colin R.
: APPLICANT: Gruenberg, Allen
: APPLICANT: Carmichael, Ireland E.
: TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: NY
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/336,345
: FILING DATE:
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Gordon, Jennifer
: REGISTRATION NUMBER: 30753
: REFERENCE/DOCKET NUMBER: 7937-006
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-8864/9741
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5049 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Parvovirus
:
: US-08-336-345-2
:
: Query Match 76.0%; Score 15.2; DB 1; Length 5049;
: Best Local Similarity 85.0%; Pred. No. 1.6e+02;
: Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
:
: Oy 1 TTAGTCTTAGTATTAGT 20
: ||||| || ||||| |||||
: Db 4745 TTAGATCATGTTAGTACTT 4764
:
: RESULT 14
: US-08-647-655-2
: Sequence 2, Application US/08647655
: Patent No. 5885585
: GENERAL INFORMATION:
: APPLICANT: Parrish, Colin R.
: APPLICANT: Gruenberg, Allen
: APPLICANT: Carmichael, Ireland E.
: TITLE OF INVENTION: Attenuated Canine Parvovirus Vaccine
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/647,655

```

```

: FILING DATE: On Even Date Herewith
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Gordon, Jennifer
: REGISTRATION NUMBER: 30,753
: REFERENCE/DOCKET NUMBER: 7937-008
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-8864/9741
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5049 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: both
: TOPOLOGY: unknown
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Parvovirus
:
: US-08-647-655-2
:
: Query Match 76.0%; Score 15.2; DB 2; Length 5049;
: Best Local Similarity 85.0%; Pred. No. 1.6e+02;
: Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
:
: Oy 1 TTAGTCTTAGTATTAGT 20
: ||||| || ||||| |||||
: Db 4745 TTAGATCATGTTAGTACTT 4764
:
: RESULT 15
: US-09-178-973B-17
: Sequence 17, Application US/09178973B
: Patent No. 6274710
: GENERAL INFORMATION:
: APPLICANT: Dumoulier, Laure
: APPLICANT: Louhed, Jamila
: APPLICANT: Renaud, Jean-Christophe
: TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible
: TITLE OF INVENTION: (TIRs)
: FILE REFERENCE: LOD 5543
: CURRENT APPLICATION NUMBER: US/09/178,973B
: CURRENT FILING DATE: 1998-10-26
: NUMBER OF SEQ ID NOS: 17
: SEQ ID NO 17
: LENGTH: 5935
: TYPE: DNA
: ORGANISM: Mus musculus
:
: US-09-178-973B-17
:
: Query Match 76.0%; Score 15.2; DB 4; Length 5935;
: Best Local Similarity 85.0%; Pred. No. 1.6e+02;
: Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
:
: Oy 1 TTAGTCTTAGTATTAGT 20
: ||||| || ||||| |||||
: Db 2996 TCATTCTTAGTATTATT 3015
:
: Search completed: April 6, 2003, 11:55:37
: Job time : 79.25 secs

```

THIS PAGE BLANK (USPTO)


```
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4669
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4669
```

```
Query Match      80.0%; Score 16; DB 9; Length 2000;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 TTAGTCTTAGTTATT 16
Db      1304 TTAGTCTTAGTTATT 1319
```

```
RESULT 3
US-09-946-807-941
; Sequence 941, Application US/09946807
; Patent No. US20020165144A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/946,807
; CURRENT FILING DATE: 2001-09-05
; PRIOR APPLICATION NUMBER: US/09/795,668
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 941
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-946-807-941
```

```
Query Match      79.0%; Score 15.8; DB 9; Length 401;
Best Local Similarity 89.5%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      1 TTAGTCTTAGTTATT 19
Db      254 TTAGTCTTAGTTATTAGT 272
```

```
RESULT 4
US-09-795-668-941
; Sequence 941, Application US/09795668
; Patent No. US20020045577A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2004-001
; CURRENT APPLICATION NUMBER: US/09/795,668
; CURRENT FILING DATE: 2001-02-28
```

```
; PRIOR APPLICATION NUMBER: US 09/515,716
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 941
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-668-941
```

```
Query Match      79.0%; Score 15.8; DB 10; Length 401;
Best Local Similarity 89.5%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      1 TTAGTCTTAGTTATTAGT 19
Db      254 TTAGTCTTAGTTATTAGT 272
```

```
RESULT 5
US-09-795-686-941
; Sequence 941, Application US/09795686
; Patent No. US20020094954A1
; GENERAL INFORMATION:
; APPLICANT: Stefansson, Hreinn
; APPLICANT: Steinthorsdottir, Valgerdur
; APPLICANT: Gulcher, Jeffrey R.
; TITLE OF INVENTION: HUMAN SCHIZOPHRENIA GENE
; FILE REFERENCE: 2345.2005-001
; CURRENT APPLICATION NUMBER: US/09/795,686
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 09/515,715
; PRIOR FILING DATE: 2000-02-28
; NUMBER OF SEQ ID NOS: 1531
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 941
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-795-686-941
```

```
Query Match      79.0%; Score 15.8; DB 10; Length 401;
Best Local Similarity 89.5%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      1 TTAGTCTTAGTTATTAGT 19
Db      254 TTAGTCTTAGTTATTAGT 272
```

```
RESULT 6
US-09-954-531-424/c
; Sequence 424, Application US/09954531
; Patent No. US20020165180A1
; GENERAL INFORMATION:
; APPLICANT: Weaver, Zoe
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using C
; FILE REFERENCE: 689290-77
; CURRENT APPLICATION NUMBER: US/09/954,531
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: US/60/233,133
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,009
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,034
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,509
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: US/60/234,567
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 1392
; SOFTWARE: PatentIn version 3.0
```

```
; SEQ ID NO 424
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)..(421)
; OTHER INFORMATION: n=a,t,g or c
US-09-954-531-424
```

```
Query Match          79.0%; Score 15.8; DB 9; Length 421;
Best Local Similarity 89.5%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 TTAGTTCTTAGCTATTACT 19
    |||||
Db 298 TTACTTCTTAGCTACT 280
```

```
RESULT 7
US-09-880-107-220/c
; Sequence 220, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-MO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 220
; LENGTH: 421
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 AAI02571
; NAME/KEY: unsure
; LOCATION: (1)..(421)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-220
```

```
Query Match          79.0%; Score 15.8; DB 10; Length 421;
Best Local Similarity 89.5%; Pred. No. 5.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 TTAGTTCTTAGCTATTACT 19
    |||||
Db 298 TTACTTCTTAGCTACT 280
```

```
RESULT 8
US-10-091-504-1359
; Sequence 1359, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1359
; LENGTH: 492
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-504-1359
```

```
Query Match          79.0%; Score 15.8; DB 9; Length 492;
Best Local Similarity 89.5%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 TTAGTTCTTAGCTATTACT 19
    |||||
Db 333 TTAGTTCTTAGCTATTACT 351
```

```
RESULT 9
US-10-091-504-1360
; Sequence 1360, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1360
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-504-1360
```

```
Query Match          79.0%; Score 15.8; DB 9; Length 492;
Best Local Similarity 89.5%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 TTAGTTCTTAGCTATTACT 19
    |||||
Db 333 TTAGTTCTTAGCTATTACT 351
```

```
RESULT 10
US-10-091-504-1361
; Sequence 1361, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1361
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-504-1361
```

```
Query Match          79.0%; Score 15.8; DB 9; Length 492;
Best Local Similarity 89.5%; Pred. No. 5.8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 TTAGTTCTTAGCTATTACT 19
    |||||
Db 333 TTAGTTCTTAGCTATTACT 351
```

```
RESULT 11
US-09-764-869-1359
; Sequence 1359, Application US/09764869
; Patent No. US20020061521A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1359
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-869-1359

Query Match
Best Local Similarity 79.0%; Score 15.8; DB 10; Length 492;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTATTAGT 19
   |||||  |||||  |||||
Db 333 TTAGTCTTAGCTATTAA 351

RESULT 12
US-09-764-869-1360
; Sequence 1360, Application US/09/764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1360
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-869-1360

Query Match
Best Local Similarity 79.0%; Score 15.8; DB 10; Length 492;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTATTAGT 19
   |||||  |||||  |||||
Db 333 TTAGTCTTAGCTATTAA 351

RESULT 13
US-09-764-869-1361
; Sequence 1361, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1361
; LENGTH: 492
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-764-869-1361

Query Match
Best Local Similarity 79.0%; Score 15.8; DB 10; Length 492;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTATTAGT 19
   |||||  |||||  |||||
Db 333 TTAGTCTTAGCTATTAA 351

RESULT 14
US-08-781-986A-994/C
; Sequence 994, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,986A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Benson, Bob
; REGISTRATION NUMBER: 30,446
; REFERENCE/DOCKET NUMBER: PB248PP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 994:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 551 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; US-08-781-986A-994

Query Match
Best Local Similarity 79.0%; Score 15.8; DB 7; Length 551;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTATTAGT 19
   |||||  |||||  |||||
Db 476 TTAGTATTGTTATTAGT 458

RESULT 15
US-09-232-785-330
; Sequence 330, Application US/09232785
; Publication No. US20030049612A1
; GENERAL INFORMATION:
; APPLICANT: International Paper Co.
; APPLICANT: Ech, Craig, S
; APPLICANT: Nelson, C. Dana
; TITLE OF INVENTION: MICROSATELLITE DNA MARKERS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 4481/1E18051
; CURRENT APPLICATION NUMBER: US/09/232,785
; CURRENT FILING DATE: 1999-01-19
; PRIOR APPLICATION NUMBER: 09/232,884
; PRIOR FILING DATE: 1999-01-15

```

```

; NUMBER OF SEQ ID NOS: 397
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 330
; LENGTH: 573
; TYPE: DNA
; ORGANISM: Pinus taeda L.
; FEATURE:
; NAME/KEY: unsure
; LOCATION: n at 319.
; OTHER INFORMATION: n is a or g or c or t/u, unknown or other.
US-09-232-785-330

```

```

Query Match          79.0%; Score 15.8; DB 9; Length 573;
Best Local Similarity 89.5%; Pred. No. 5.9e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 TTAGTCTTACTTATTAGT 19
    ||||||||| 1 |||||||
Db 196 TTAGTCTTGTGATATTAGT 214

```

Search completed: April 6, 2003, 12:09:29
 Job time : 60 secs

THIS PAGE BLANK (USPTO)

Lehmann, R., Baumgart, C., Parra, G., April, J. F., Guigo, R., Kump, K., Cox, E., Quail, M. A., Platzer, M., Rosenthal, A., and Tunngdal, B.

TITLE
JOURNAL
REMARK
REFERENCE
AUTHORS
JOURNAL
COMMENT

Noegel, A.A.
Sequence and Analysis of Chromosome 2 of Dictyostelium
Unpublished
The Dictyostelium Genome Sequencing Consortium
2 (bases 1 to 8659)
Baumgart, C.
Direct Submission
Submitted (09-APR-2002) Genome Analysis, Institute of Molecular
Biotechnology, Beutenbergstr. 11, Jena 07745, Germany
CDS predictions from Genedit may contain errors. Further information
is available from IMB Jena, Department of Genome Analysis
(http://genome.imb-jena.de/dictyostelium/)
and the University Cologne, Institute for Biochemistry I
(http://www.uni-koeln.de/dictyostelium/project.shtml)
Funding
Agency : Deutsche Forschungsgemeinschaft (DFG).
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1. 8659
/organism="Dictyostelium discoideum"
/strain="AX4"
/db_xref="taxon:44689"
/chromosome="2"
/map="3648613-3657270"
BASE COUNT 3236 a 1121 c 1242 g 3060 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 8659;
Best Local Similarity 95.0%; Pred. No. 29;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTACTCTAGTTTACTT 20
||||| |||||||||
Db 577 TTACTTTTACTTTTACTT 558

RESULT 3
AC108390
LOCUS
DEFINITION Homo sapiens chromosome 8 clone RP11-2342N23 map 8, LOW-PASS
SEQUENCE SAMPLING.
AC108390
VERSION AC108390.1 GI:18377195
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 63673)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 8, clone RP11-2342N23
Unpublished
2 (bases 1 to 63673)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouhgalter, B.,
Brown, A., Camarata, J., Campolano, A., Chang, J., Chazaro, B.,
Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., DeArrelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Grinde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karacas, A., Kells, C., Labrecque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C.,
Macdonald, P., Major, J., Margis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., Meldim, J., Menus, L., Mihova, T.,
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Riebeck, M., Riley, R., Rise, C., Rogov, P., Roman, J.,
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S.,

TITLE
JOURNAL
COMMENT

Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wymen, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (27-JAN-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: LI3063
Center clone name: 2342_N_23

* NOTE: This record contains 83 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
666 765: gap of 100 bp
766 1416: contig of 651 bp in length
1417 1516: gap of 100 bp
1517 2160: contig of 644 bp in length
2161 2260: gap of 100 bp
2261 2912: contig of 652 bp in length
2913 3012: gap of 100 bp
3013 3681: contig of 669 bp in length
3682 3781: gap of 100 bp
3782 4420: contig of 639 bp in length
4421 4520: gap of 100 bp
4521 5167: contig of 647 bp in length
5168 5267: gap of 100 bp
5268 5922: contig of 655 bp in length
5923 6022: gap of 100 bp
6023 6676: contig of 654 bp in length
6677 6776: gap of 100 bp
6777 7453: contig of 677 bp in length
7454 7553: gap of 100 bp
7554 8255: contig of 702 bp in length
8256 8355: gap of 100 bp
8356 9044: contig of 689 bp in length
9045 9144: gap of 100 bp
9145 9805: contig of 661 bp in length
9806 9905: gap of 100 bp
9906 10548: contig of 643 bp in length
10549 10648: gap of 100 bp
10649 11310: contig of 662 bp in length
11311 11410: gap of 100 bp
11411 12087: contig of 677 bp in length
12088 12187: gap of 100 bp
12188 12872: contig of 685 bp in length
12873 12972: gap of 100 bp
12973 13629: contig of 657 bp in length
13630 13729: gap of 100 bp
13730 14406: contig of 677 bp in length
14407 14506: gap of 100 bp
14507 15177: contig of 671 bp in length
15178 15277: gap of 100 bp
15278 15930: contig of 653 bp in length
15931 16030: gap of 100 bp


```

* 16031 16687: contig of 657 bp in length
* 16688 16787: gap of 100 bp
* 16788 17450: contig of 663 bp in length
* 17451 17550: gap of 100 bp
* 17551 18222: contig of 672 bp in length
* 18223 18322: gap of 100 bp
* 18323 18978: contig of 656 bp in length
* 18979 19078: gap of 100 bp
* 19079 19751: contig of 673 bp in length
* 19752 19851: gap of 100 bp
* 19852 20523: contig of 672 bp in length
* 20524 20623: gap of 100 bp
* 20624 21313: contig of 690 bp in length
* 21314 21413: gap of 100 bp
* 21414 22091: contig of 678 bp in length
* 22092 22191: gap of 100 bp
* 22192 22879: contig of 688 bp in length
* 22880 22979: gap of 100 bp
* 22980 23636: contig of 657 bp in length
* 23637 23736: gap of 100 bp
* 23737 24426: contig of 690 bp in length
* 24427 24526: gap of 100 bp
* 24527 25215: contig of 689 bp in length
* 25216 25315: gap of 100 bp
* 25316 25971: contig of 656 bp in length
* 25972 26071: gap of 100 bp
* 26072 26730: contig of 659 bp in length
* 26731 26830: gap of 100 bp
* 26831 27503: contig of 673 bp in length
* 27504 27603: gap of 100 bp
* 27604 28282: contig of 679 bp in length
* 28283 28382: gap of 100 bp
* 28383 29059: contig of 677 bp in length
* 29060 29159: gap of 100 bp
* 29160 29841: contig of 682 bp in length
* 29842 29941: gap of 100 bp
* 29942 30594: contig of 653 bp in length
* 30595 30694: gap of 100 bp
* 30695 31374: contig of 680 bp in length
* 31375 31474: gap of 100 bp
* 31475 32130: contig of 656 bp in length
* 32131 32230: gap of 100 bp
* 32231 32886: contig of 656 bp in length
* 32887 32986: gap of 100 bp
* 32987 33654: contig of 668 bp in length
* 33655 33754: gap of 100 bp
* 33755 34411: contig of 657 bp in length
* 34412 34511: gap of 100 bp
* 34512 35169: contig of 658 bp in length
* 35170 35269: gap of 100 bp
* 35270 35944: contig of 675 bp in length
* 35945 36044: gap of 100 bp
* 36045 36707: contig of 663 bp in length
* 36708 36807: gap of 100 bp
* 36808 37490: contig of 683 bp in length
* 37491 37590: gap of 100 bp
* 37591 38253: contig of 663 bp in length
* 38254 38353: gap of 100 bp
* 38354 39023: contig of 670 bp in length
* 39024 39123: gap of 100 bp
* 39124 39611: contig of 688 bp in length
* 39612 39911: gap of 100 bp
* 39912 40591: contig of 680 bp in length
* 40592 40691: gap of 100 bp
* 40692 41359: contig of 668 bp in length
* 41360 41459: gap of 100 bp
* 41460 42122: contig of 663 bp in length
* 42123 42222: gap of 100 bp
* 42223 42898: contig of 676 bp in length
* 42899 42998: gap of 100 bp
* 42999 43685: contig of 687 bp in length
* 43686 43785: gap of 100 bp
* 43786 44450: contig of 665 bp in length

```

```

* 44451 44550: gap of 100 bp
* 44551 45203: contig of 653 bp in length
* 45204 45303: gap of 100 bp
* 45304 45976: contig of 673 bp in length
* 45977 46076: gap of 100 bp
* 46077 46718: contig of 642 bp in length
* 46719 46818: gap of 100 bp
* 46819 47499: contig of 681 bp in length
* 47500 47599: gap of 100 bp
* 47600 48268: contig of 669 bp in length
* 48269 48368: gap of 100 bp
* 48369 49028: contig of 660 bp in length
* 49029 49128: gap of 100 bp
* 49129 49781: contig of 653 bp in length
* 49782 49881: gap of 100 bp
* 49882 50555: contig of 674 bp in length
* 50556 50655: gap of 100 bp
* 50656 51315: contig of 660 bp in length
* 51316 51415: gap of 100 bp
* 51416 52088: contig of 673 bp in length
* 52089 52188: gap of 100 bp
* 52189 52860: contig of 672 bp in length

```

```

Query Match          92.0%: Score 18.4; DB 1; Length 63673;
Best Local Similarity 95.0%: Pred. No. 36;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTACTTTTACTT 20
Db 44734 TTAGTTTGTAGTTTACTT 44753

```

```

RESULT 4
AC015208/c 64195 bp DNA linear HTG 16-NOV-1999
DEFINITION
Drosophila melanogaster. *** SEQUENCING IN PROGRESS ***; In ordered
pieces.
ACCESSION
AC015208
VERSION
AC015208.1 GI:6436127
KEYWORDS
HTG; HTGS_PHASE2.
SOURCE
Drosophila melanogaster.
ORGANISM
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 64195)
REFERENCE
Adams, M. and Venter, J.C.
AUTHORS
Direct Submission
TITLE
Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT
This sequence was identified as CDM:10211252 by the submitter.
For further information on this sequence e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

```

```

FEATURES
source
1..64195
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

```

```

BASE COUNT 17633 a 14801 c 14632 g 17129 t
ORIGIN

```

```

Query Match          92.0%: Score 18.4; DB 1; Length 64195;
Best Local Similarity 95.0%: Pred. No. 36;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTACTTTTACTT 20

```

```

Db 4393 TTAGTCATAGTTTACTT 4374

```

```

RESULT 5

```

AC121123 70307 bp DNA linear HTG 15-MAY-2002
LOCUS Mus musculus clone RP23-474E16, LOW-PASS SEQUENCE SAMPLING.
DEFINITION AC121123
ACCESSION AC121123
VERSION AC121123.1 GI:20800260
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 70307)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
JOURNAL Mus musculus, clone RP23-474E16
REFERENCE 2 (bases 1 to 70307)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campolano, A., Chang, J.,
Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazares, R.,
Lander, S., Lehoczkay, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (15-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smith, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L22814
Center clone name: 474_E_16

* NOTE: This record contains 87 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.
* 1 702: contig of 702 bp in length
* 703 802: gap of 100 bp
* 803 1523: contig of 721 bp in length
* 1524 1623: gap of 100 bp
* 1624 2341: contig of 718 bp in length
* 2342 2441: gap of 100 bp
* 2442 3142: contig of 701 bp in length
* 3143 3242: gap of 100 bp
* 3243 3945: contig of 703 bp in length
* 3946 4045: gap of 100 bp
* 4046 4737: contig of 692 bp in length
* 4738 4837: gap of 100 bp
* 4838 5501: contig of 664 bp in length
* 5502 5601: gap of 100 bp
* 5602 6315: contig of 714 bp in length
* 6316 6415: gap of 100 bp
* 6416 7138: contig of 723 bp in length
* 7139 7238: gap of 100 bp
* 7239 7964: contig of 726 bp in length
* 7965 8064: gap of 100 bp
* 8065 8765: contig of 701 bp in length
* 8766 8865: gap of 100 bp
* 8866 9578: contig of 713 bp in length
* 9579 9678: gap of 100 bp
* 9679 10404: contig of 726 bp in length
* 10405 11206: contig of 702 bp in length
* 11207 11306: gap of 100 bp
* 11307 12002: contig of 696 bp in length
* 12003 12102: gap of 100 bp
* 12103 12806: contig of 704 bp in length
* 12807 12906: gap of 100 bp
* 12907 13614: contig of 708 bp in length
* 13615 13714: gap of 100 bp
* 13715 14426: contig of 712 bp in length
* 14427 14526: gap of 100 bp
* 14527 15253: contig of 727 bp in length
* 15254 15353: gap of 100 bp
* 15354 16083: contig of 730 bp in length
* 16084 16183: gap of 100 bp
* 16184 16877: contig of 694 bp in length
* 16878 16977: gap of 100 bp
* 16978 17694: contig of 717 bp in length
* 17695 17794: gap of 100 bp
* 17795 18510: contig of 716 bp in length
* 18511 18610: gap of 100 bp
* 18611 19309: contig of 699 bp in length
* 19310 19409: gap of 100 bp
* 19410 20118: contig of 709 bp in length
* 20119 20218: gap of 100 bp
* 20219 20909: contig of 691 bp in length
* 20910 21009: gap of 100 bp
* 21010 21702: contig of 693 bp in length
* 21703 21802: gap of 100 bp
* 21803 22526: contig of 724 bp in length
* 22527 22626: gap of 100 bp
* 22627 23333: contig of 707 bp in length
* 23334 23433: gap of 100 bp
* 23434 24154: contig of 721 bp in length
* 24155 24254: gap of 100 bp
* 24255 24960: contig of 726 bp in length
* 24961 25080: gap of 100 bp
* 25081 25785: contig of 705 bp in length
* 25786 25885: gap of 100 bp
* 25886 26609: contig of 724 bp in length
* 26610 26709: gap of 100 bp
* 26710 27439: contig of 730 bp in length
* 27440 27539: gap of 100 bp
* 27540 28243: contig of 704 bp in length
* 28244 28343: gap of 100 bp
* 28344 29032: contig of 689 bp in length
* 29033 29132: gap of 100 bp
* 29133 29816: contig of 684 bp in length
* 29817 29916: gap of 100 bp
* 29917 30611: contig of 695 bp in length
* 30612 30711: gap of 100 bp
* 30712 31407: contig of 696 bp in length
* 31408 31507: gap of 100 bp
* 31508 32234: contig of 727 bp in length
* 32235 32334: gap of 100 bp
* 32335 33066: contig of 732 bp in length
* 33067 33166: gap of 100 bp
* 33167 33867: contig of 701 bp in length

```

* 3368 33967: gap of 100 bp
* 33968 34682: contig of 715 bp in length
* 34683 34782: gap of 100 bp
* 34783 35508: contig of 726 bp in length
* 35509 35608: gap of 100 bp
* 35609 36286: contig of 678 bp in length
* 36287 36386: gap of 100 bp
* 36387 37094: contig of 708 bp in length
* 37095 37194: gap of 100 bp
* 37195 37885: contig of 691 bp in length
* 37886 37985: gap of 100 bp
* 37986 38687: contig of 702 bp in length
* 38688 38787: gap of 100 bp
* 38788 39487: contig of 700 bp in length
* 39488 39587: gap of 100 bp
* 39588 40290: contig of 703 bp in length
* 40291 40390: gap of 100 bp
* 40391 41120: contig of 730 bp in length
* 41121 41220: gap of 100 bp
* 41221 41948: contig of 728 bp in length
* 41949 42048: gap of 100 bp
* 42049 42756: contig of 708 bp in length
* 42757 42856: gap of 100 bp
* 42857 43581: contig of 725 bp in length
* 43582 43681: gap of 100 bp
* 43682 44406: contig of 725 bp in length
* 44407 44506: gap of 100 bp
* 44507 45207: contig of 701 bp in length
* 45208 45307: gap of 100 bp
* 45308 46004: contig of 697 bp in length
* 46005 46104: gap of 100 bp
* 46105 46787: contig of 683 bp in length
* 46788 46887: gap of 100 bp
* 46888 47572: contig of 685 bp in length
* 47573 47672: gap of 100 bp
* 47673 48391: contig of 719 bp in length
* 48392 48491: gap of 100 bp
* 48492 49198: contig of 707 bp in length
* 49199 49298: gap of 100 bp
* 49299 50024: contig of 726 bp in length
* 50025 50124: gap of 100 bp
* 50125 50857: contig of 733 bp in length
* 50858 50957: gap of 100 bp
* 50958 51684: contig of 727 bp in length
* 51685 51784: gap of 100 bp
* 51785 52487: contig of 703 bp in length
* 52488 52587: gap of 100 bp
* 52588 53313: contig of 726 bp in length
* 53314 53413: gap of 100 bp
* 53414 54132: contig of 719 bp in length
* 54133 54232: gap of 100 bp
* 54233 54949: contig of 717 bp in length
* 54950 55049: gap of 100 bp
* 55050 55726: contig of 677 bp in length
* 55727 55826: gap of 100 bp

```

Query Match 92.0%: Score 18.4; DB 1; Length 70307;
 Best Local Similarity 95.0%: Pred. No. 36;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTCTCTAGTTTGTAGTT 20
 Db 20430 TCAGTCTCTAGTTTGTAGTT 20449

RESULT 6
 AC11421/c 77961 bp DNA HTG 13-JUL-2002
 LOCUS Rattus norvegicus clone CH230-45G18, *** SEQUENCING IN PROGRESS
 DEFINITION *** 44 unordered pieces.
 ACCESSION AC11421
 KEYWORDS AC11421.2 GI:21735962
 HTG: HTGS_PHASE1.

SOURCE
 ORGANISM
 Norway rat.
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 AUTHORS
 1 (bases 1 to 77961)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amarantunge,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbara,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Day-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
 Homsi,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., Kling,L., Korvah,J., Kovar,C.,
 Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H.,
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokwenko,S., Oguh,M., Okunogun,G.,
 Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojudoakan,I., Rolfe,M., Ruiz,S., Severy,G.,
 Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,
 Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansley,J., Taylor,C., Taylor,T., Telitod,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalton,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
 Williams,G., Williamson,A., Wlecezyk,R., Woodson,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G. and Gibbs,R.

TITLE
 JOURNAL
 Direct Submission
 Unpublished
 2 (bases 1 to 77961)
 REFERENCE
 AUTHORS
 Worley,K.C.
 TITLE
 JOURNAL
 Direct Submission
 Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 77961)
 REFERENCE
 AUTHORS
 Worley,K.C.
 TITLE
 JOURNAL
 Direct Submission
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 12, 2002 this sequence version replaced gi:18701185.
 COMMENT
 Genome Center

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 Project Information
 Center project name: GMG
 Center Clone name: CH230-45G18
 Summary Statistics
 Sequencing vector: Plasmid
 Chemistry: Dye-terminator Big Dye 100% of reads
 Assembly program: Phrap; version 0.990329

Consensus quality: 37596 bases at least Q40
Consensus quality: 39809 bases at least Q30
Consensus quality: 41410 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.ngsc.bcm.tmc.edu/docs/gendank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 44 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1012: contig of 1012 bp in length
* 1013 1112: gap of unknown length
* 1113 2193: contig of 1081 bp in length
* 2194 2293: gap of unknown length
* 2294 3348: contig of 1055 bp in length
* 3349 4839: gap of unknown length
* 4840 4939: contig of 1391 bp in length
* 4940 6237: gap of unknown length
* 6238 6337: contig of 1298 bp in length
* 6338 7393: gap of unknown length
* 7394 7493: contig of 1056 bp in length
* 7494 8675: gap of unknown length
* 8676 8775: contig of 1182 bp in length
* 8776 10195: gap of unknown length
* 10196 10295: contig of 1420 bp in length
* 10296 11995: gap of unknown length
* 11996 12095: contig of 1700 bp in length
* 12096 13297: gap of unknown length
* 13298 13397: contig of 1202 bp in length
* 13398 14707: gap of unknown length
* 14708 14807: contig of 1310 bp in length
* 14808 16094: gap of unknown length
* 16095 16194: contig of 1287 bp in length
* 16195 17588: gap of unknown length
* 17589 17689: contig of 1394 bp in length
* 17690 18894: gap of unknown length
* 18895 18994: contig of 1206 bp in length
* 18995 20450: gap of unknown length
* 20451 20550: contig of 1456 bp in length
* 20551 21565: gap of unknown length
* 21566 21665: contig of 1015 bp in length
* 21666 23430: gap of unknown length
* 23431 23530: contig of 1765 bp in length
* 23531 25263: gap of unknown length
* 25264 25363: contig of 1733 bp in length
* 25364 26692: gap of unknown length
* 26693 26792: contig of 1329 bp in length
* 26793 28337: gap of unknown length
* 28338 28437: contig of 1545 bp in length
* 28439 29787: gap of unknown length
* 29788 29887: contig of 1350 bp in length
* 29888 32005: gap of unknown length
* 32006 32105: contig of 2118 bp in length
* 32106 33963: gap of unknown length
* 33964 34063: contig of 1858 bp in length
* 34064 35374: gap of unknown length
* 35375 35474: contig of 1311 bp in length
* 35475 36824: gap of unknown length
* 36825 36924: contig of 1350 bp in length
* 36926 38528: gap of unknown length
* 38529 38628: contig of 1604 bp in length
* 38629 40883: gap of unknown length
* 40884 40983: contig of 2255 bp in length
* 40984 42742: gap of unknown length
* 42743 42842: contig of 1759 bp in length
* 42843 44303: gap of unknown length
* 44304 44403: contig of 1461 bp in length
* 44404 46331: gap of unknown length
* 46331: contig of 1928 bp in length

46332 46431: gap of unknown length
* 46432 48139: contig of 1708 bp in length
* 48140 48239: gap of unknown length
* 48240 49958: contig of 1719 bp in length
* 49959 50059: gap of unknown length
* 50060 51852: contig of 1794 bp in length
* 51853 51953: gap of unknown length
* 51954 53574: contig of 1622 bp in length
* 53575 53675: gap of unknown length
* 53676 55350: contig of 1675 bp in length
* 55351 55449: gap of unknown length
* 55450 56989: contig of 1540 bp in length
* 56990 57089: gap of unknown length
* 57090 58922: contig of 1833 bp in length
* 58923 59022: gap of unknown length
* 59023 61257: contig of 2235 bp in length
* 61258 61357: gap of unknown length
* 61358 63981: contig of 2624 bp in length
* 63982 64081: gap of unknown length
* 64082 66168: contig of 2087 bp in length
* 66169 66268: gap of unknown length
* 66269 68952: contig of 2684 bp in length
* 68953 69052: gap of unknown length
* 69053 71673: contig of 2620 bp in length
* 71674 71772: gap of unknown length
* 71773 75048: contig of 3276 bp in length
* 75049 75149: gap of unknown length
* 75149 77961: contig of 2813 bp in length.
location/Qualifiers
1. 77961
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-45G18"
BASE COUNT 21320 a 14828 c 14980 g 21355 t 5478 others
ORIGIN
Query Match 92.0% Score 18.4: DB 1: Length 77961;
Best Local Similarity 95.0% Pred. No. 36;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 TTAGTCTTACTTTTACTT 20
Db 16419 TAACTTCTTACTTTTACTT 16400
RESULT 7 104887 bp DNA 11near HTG 06-AUG-2002
AC104044
LOCUS
DEFINITION
HOMO sapiens chromosome 8 clone CTD-2219A7 map 8, *** SEQUENCING IN
PROGRESS ***, 2 ordered pieces.
AC104044
AC104044.4 GI:22123469
VERSION
HTG: HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 104887)
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Birn, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouhgalter, B.,
Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N.,
Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Labocque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,

McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrum, J.,
Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunhkanh, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Roselli, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggiani, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS
Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 104887)

Barron, B., Nushbaum, C., Lander, E., All, A., Allen, N., Anderson, S.,
Barra, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
Camata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A.,
Karats, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Lin, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunhkanh, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT
Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 6, 2002 this sequence version replaced g1:18643594.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
Project Information
Center project name: 2219_A_7

FEATURES
source
* NOTE: This is a 'working draft' sequence. It currently
* consists of 2 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 27562: contig of 27562 bp in length
* 27563 27562: gap of 100 bp
* 27563 104887: contig of 77225 bp in length.
Location/Qualifiers
1..104887

BASE COUNT
ORIGIN
Query Match
31026 a 20095 c 20706 g 32141 t 915 others
92.0%; Score 18.4; DB 1; length 104887;

Best Local Similarity 95.0%; Pred. No. 38;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTACTTTTACTT 20
Db 35479 TTAGTCTTACTTTTACTT 35498

RESULT 8
AC005504/C

LOCUS
DEFINITION
AC005504
Plasmodium falciparum chromosome 12, *** SEQUENCING IN PROGRESS
***, 3 unordered pieces.

AC005504.3 GI:4558584
HTG: HTGS_PHASE1
KEYWORDS
SOURCE
Plasmodium falciparum.
Plasmodium falciparum.
Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE
AUTHORS
Hyman, R.W., Fung, E.L., Qin, F., Tamaki, T., Kurd, O.B., Conway, A.B.
and Davis, R.W.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Submitted (21-AUG-1998) Stanford DNA Sequencing and Technology
Center, Stanford University, 855 California Avenue, Palo Alto, CA
94304, USA

COMMENT

On Apr 2, 1999 this sequence version replaced g1:4337172.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 58642: contig of 58642 bp in length
* 58643 58642: gap of unknown length
* 58843 91011: contig of 32169 bp in length
* 91012 91211: gap of unknown length
* 91212 104992: contig of 13781 bp in length.

FEATURES
source
Location/Qualifiers
1..104992

/organism="Plasmodium falciparum"
/db_xref="taxon:5833"
/chromosome="12"

BASE COUNT
ORIGIN
44286 a 9326 c 9564 g 41411 t 405 others

Query Match
Best Local Similarity 95.0%; Pred. No. 38;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTACTTTTACTT 20
Db 57851 TTAGTCTTACTTTTACTT 57832

RESULT 9
AC004710/C

LOCUS
DEFINITION
AC004710
Plasmodium falciparum chromosome 12, *** SEQUENCING IN PROGRESS
***, 2 unordered pieces.

AC004710.3 GI:4558577
HTG: HTGS_PHASE1
KEYWORDS
SOURCE
Plasmodium falciparum.
Plasmodium falciparum.
Plasmodium falciparum.
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE
1 (bases 1 to 114226)

AUTHORS Hyman,R.W., Fung,E.L., Qin,F., Tamaki,T., Kurd,I.O.B., Conway,A.B. and Davis,R.W.

JOURNAL Plasmodium falciparum 3D7 chromosome 12

REFERENCE 2 (bases 1 to 114226)

AUTHORS Hyman,R.W., Qin,F., Fung,E.L., Conway,A.B. and Davis,R.W.

JOURNAL Direct Submission

Submitted (21-MAY-1998) Stanford DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

COMMENT On Apr 2, 1999 this sequence version replaced gi:4337166.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved

1 93069: contig of 93069 bp in length

* 93070 93269: gap of unknown length

* 93270 114226: contig of 20957 bp in length.

location/Qualifiers

1. 114226

/organism="Plasmodium falciparum"

/db_xref="taxon:5833"

/chromosome="12"

BASE COUNT 44338 a 11954 c 10111 g 47622 t 201 others

ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 114226;

Best Local Similarity 95.0%; Pred. No. 38;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTACTCTTACTTTTACTT 20

||||| |||||||||

Db 92278 TTACTTTTACTTTTACTT 92259

RESULT 10

AC002421 131347 bp DNA linear HTG 13-JUN-2002

LOCUS Homo sapiens chromosome X clone PMXD1, *** SEQUENCING IN PROGRESS

DEFINITION *** 4 unordered pieces.

AC002421

VERSION AC002421.2 GI:21405641

KEYWORDS HTG; HTGS_PHASE1.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 131347)

Chen,E., Brownstein,B.H., States,D.J., Schlessinger,D. and Mazzarella,R.

Direct Submission

Unpublished (1997)

2 (bases 1 to 131347)

Brownstein,B.H., States,D.J. and Mazzarella,R.

Submitted (12-AUG-1997) Center for Genetics in Medicine, Box 8232, Washington University School of Medicine, 4566 Scott Avenue, St. Louis, MO 63110, USA

On Jun 13, 2002 this sequence version replaced gi:2323248.

Current status of this project is available at:

'http://genome.wustl.edu/cgm/seq_projects.html'

Submitted by:

Elison Chen,

Advanced Center for Genetic Technology,

Applied Biosystems Division of Perlin Elmer Corp.,

850 Lincoln Center Drive,

Foster City, CA 94404 USA

e-mail: ellison@genseq.apidbio.com

and

Buddy Brownstein,

Center for Genetics in Medicine,

Washington University School of Medicine, Box 8232

4566 Scott Avenue

St. Louis, MO 63110, USA

e-mail: buddy@genetics.wustl.edu

and

David J. States,

Institute for Biomedical Computing

Washington University in St. Louis

700 South Euclid Ave.

St. Louis, MO 63108 USA

e-mail: states@bioc.wustl.edu.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 4 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved

1 22595: contig of 22595 bp in length

* 22596 22695: gap of 100 bp

* 22696 66983: contig of 44288 bp in length

* 66984 67083: gap of 100 bp

* 67084 122847: contig of 55764 bp in length

* 122848 122947: gap of 100 bp

* 122948 131347: contig of 8400 bp in length.

location/Qualifiers

1. 131347

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="X"

/clone="PMXD1"

BASE COUNT 40590 a 23040 c 23923 g 43494 t 300 others

ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 131347;

Best Local Similarity 95.0%; Pred. No. 39;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTACTTTTACTT 20

||||| |||||||||

Db 45394 TTATTTCTTACTTTTACTT 45413

RESULT 11

AC130805 139205 bp DNA linear HTG 21-AUG-2002

LOCUS Medicago truncatula clone mth2-9n1, WORKING DRAFT SEQUENCE, 7

DEFINITION unordered pieces.

AC130805

VERSION AC130805.2 GI:22325011

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE barrel medic.

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

1 (bases 1 to 139205)

Shauli,S., Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.

Medicago truncatula BAC Clone mth2-9n1

Unpublished

2 (bases 1 to 139205)

Shauli,S., Lin,S., Dixon,R., May,G., Summer,L., Gonzales,B., Cook,D., Kim,D. and Roe,B.A.

Direct Submission

JOURNAL Submitted (14-AUG-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA

REFERENCE 3 (bases 1 to 139205)

AUTHORS Shaull, S., Lin, S., Dixon, R., May, G., Sumner, L., Gonzales, B., Cook, D., Kim, D., and Roe, B. A.

TITLE Direct Submission

JOURNAL Submitted (21-AUG-2002) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA

COMMENT On Aug 20, 2002 this sequence version replaced gi:22218497.

----- Genome Center
Center: Department Of Chemistry And Biochemistry
The University Of Oklahoma
Center code: UOKNOR

* NOTE: This is a 'working draft' sequence. It currently consists of 7 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved.

1 2045: contig of 2045 bp in length
* 2046 2145: gap of unknown length
* 2146 5410: contig of 3265 bp in length
* 5411 5510: gap of unknown length
* 5511 11789: contig of 6279 bp in length
* 11790 11889: gap of unknown length
* 11890 18821: contig of 6932 bp in length
* 18822 18921: gap of unknown length
* 18922 35712: contig of 16791 bp in length
* 35713 35812: gap of unknown length
* 35813 65594: contig of 29782 bp in length
* 65595 139205: contig of 73511 bp in length.
* 65695

FEATURES
source
1. 139205
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="mtb2-9n1"
/clone_lib="Medicago truncatula BAC library H2"
BASE COUNT 47587 a 23034 c 21787 g 46196 t 601 others

ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 139205;
Best Local Similarity 95.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTACGTTTACTT 20
|||||
Db 96871 TTAGTCTTACGTTTACTT 96852

RESULT 12
AC109061
LOCUS AC109061 142423 bp DNA linear HTG 17-JUL-2002
DEFINITION Rattus norvegicus clone CH230-101L7, *** SEQUENCING IN PROGRESS
***, 69 unordered pieces.
ACCESSION AC109061
VERSION AC109061.3 GI:21744138
KEYWORDS HTG: HTGS, PHASEL.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
1 (bases 1 to 142423)
Munry, D. M., Adams, C., Adio-Oduola, B., Ali-Osman, F. R., Allen, C., Alsbrooks, S. L., Amaralunge, H. C., Are, J. R., Ayele, M., Banks, T., Barbata, J., Benton, J., Blinze, K., Blankenburg, K., Bonnin, D., Bouck, J., Bowle, S., Brieva, M., Brown, E., Brown, M., Bryant, N. P.,

Buhay, C., Burch, P., Burkett, C., Burrell, K. L., Byrd, N. C., Carron, T. P., Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C. D., Cox, C., Coyle, M. D., Dathorne, S. R., David, R., Davila, M. L., Davis, C., Davy-Carroll, J., Dedrich, D. A., Delaney, K. R., Delgado, O., Denn, A. L., Ding, Y., Dinh, H. H., Douthwaite, K. J., Draper, H., Dugan-Rocha, S., Durbin, K. J., Eamhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frenzt, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Hollway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Huiyk, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Kovach, J., Kovar, C., Kratochvic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Louised, H., Lozado, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhney, E., McLeod, M. P., Meador, M., Mel, G., Metker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I., Sodergren, E., Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Swalek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R., Wang, S., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S., Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D., Welshstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 142423)
Worley, K. C.
Direct Submission
Submitted (03-FEB-2002) Human Genome Sequencing Center, Department Of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 142423)
Worley, K. C.
Direct Submission
Submitted (17-JUL-2002) Human Genome Sequencing Center, Department Of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Jul 13, 2002 this sequence version replaced gi:18846789.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GPV
Center clone name: CH230-101L7
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 84036 bases at least Q40
Consensus quality: 88564 bases at least Q30
Consensus quality: 92346 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_drift_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 69 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1004: contig of 1004 bp in length
* 1005 1104: gap of unknown length
* 1105 2142: contig of 1038 bp in length
* 2143 2242: gap of unknown length
* 2243 3287: contig of 1045 bp in length
* 3288 3388: gap of unknown length
* 3388 4626: contig of 1239 bp in length
* 4627 4727: gap of unknown length
* 4727 5837: contig of 1111 bp in length
* 5838 5937: gap of unknown length
* 5938 7447: contig of 1510 bp in length
* 7448 7547: gap of unknown length
* 7548 8871: contig of 1324 bp in length
* 8872 8971: gap of unknown length
* 8972 10001: contig of 1030 bp in length
* 10002 10101: gap of unknown length
* 10102 11384: contig of 1283 bp in length
* 11385 11484: gap of unknown length
* 11485 12538: contig of 1054 bp in length
* 12539 12638: gap of unknown length
* 12639 13875: contig of 1237 bp in length
* 13876 13975: gap of unknown length
* 13976 15308: contig of 1533 bp in length
* 15309 15608: gap of unknown length
* 15609 16953: contig of 1345 bp in length
* 16954 17053: gap of unknown length
* 17054 18233: contig of 1180 bp in length
* 18234 18333: gap of unknown length
* 18334 19344: contig of 1011 bp in length
* 19345 19444: gap of unknown length
* 19445 20666: contig of 1222 bp in length
* 20667 20766: gap of unknown length
* 20767 22253: contig of 1487 bp in length
* 22254 22353: gap of unknown length
* 22354 23999: contig of 1646 bp in length
* 24000 24099: gap of unknown length
* 24100 25396: contig of 1297 bp in length
* 25397 25496: gap of unknown length
* 25497 27117: contig of 1621 bp in length
* 27118 27217: gap of unknown length
* 27218 28890: contig of 1673 bp in length
* 28891 28990: gap of unknown length
* 28991 31178: contig of 2188 bp in length
* 31179 31278: gap of unknown length
* 31279 32681: contig of 1403 bp in length
* 32682 32781: gap of unknown length
* 32782 34782: contig of 2001 bp in length
* 34783 34882: gap of unknown length
* 34883 35991: contig of 1109 bp in length
* 35992 36091: gap of unknown length
* 36092 37963: contig of 1872 bp in length
* 37964 38063: gap of unknown length
* 38064 40364: contig of 2301 bp in length
* 40365 40464: gap of unknown length
* 40465 41813: contig of 1349 bp in length
* 41814 41913: gap of unknown length
* 41914 43241: contig of 1328 bp in length
* 43242 43341: gap of unknown length
* 43342 44828: contig of 1487 bp in length
* 44829 44928: gap of unknown length
* 44929 46483: contig of 1555 bp in length
* 46484 46583: gap of unknown length
* 46584 47891: contig of 1308 bp in length
* 47892 47991: gap of unknown length
* 47992 49446: contig of 1455 bp in length
* 49447 49546: gap of unknown length
* 49547 51449: contig of 1903 bp in length
* 51450 53745: gap of unknown length
* 53745: contig of 2196 bp in length

53746 53845: gap of unknown length
* 53846 56029: contig of 2184 bp in length
* 56030 56129: gap of unknown length
* 56130 57995: contig of 1866 bp in length
* 57996 58095: gap of unknown length
* 58096 59787: contig of 1691 bp in length
* 59788 59886: gap of unknown length
* 59887 61264: contig of 1378 bp in length
* 61265 61364: gap of unknown length
* 61365 62623: contig of 1259 bp in length
* 62624 62723: gap of unknown length
* 62724 64854: contig of 2131 bp in length
* 64855 64954: gap of unknown length
* 64955 67347: contig of 2393 bp in length
* 67348 67447: gap of unknown length
* 67448 69033: contig of 1586 bp in length
* 69034 69133: gap of unknown length
* 69134 70383: contig of 1250 bp in length
* 70384 70483: gap of unknown length
* 70484 72010: contig of 1527 bp in length
* 72011 72110: gap of unknown length
* 72111 74481: contig of 2371 bp in length
* 74482 74581: gap of unknown length
* 74582 76837: contig of 2256 bp in length
* 76838 76937: gap of unknown length
* 76938 79589: contig of 2652 bp in length
* 79590 79689: gap of unknown length
* 79690 82043: contig of 2354 bp in length
* 82044 82143: gap of unknown length
* 82144 84899: contig of 2756 bp in length
* 84900 84999: gap of unknown length
* 85000 87010: contig of 2011 bp in length
* 87011 87110: gap of unknown length
* 87111 89409: contig of 2299 bp in length
* 89410 89509: gap of unknown length
* 89510 91888: contig of 2379 bp in length
* 91889 91988: gap of unknown length
* 91989 94785: contig of 2797 bp in length

Query Match 92.0%; Score 18.4; DB 1; Length 142423;
Best Local Similarity 95.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTTAGTT 20
Db 86537 TTAGTCTTAGTCTTTAGTT 86556
|||||
|||||

RESULT 13
AC130755 153671 bp DNA 1linear HTG 14-AUG-2002
LOCUS
DEFINITION
Rattus norvegicus clone CH230-10D6, *** SEQUENCING IN PROGRESS ***
AC130755
77 unordered pieces.
AC130755 1 GI:22218406
VERSION
HTG: HTGS_PHASE1.
KEYWORDS
SOURCE
ORGANISM
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 153671)
Muzny,D.,Marle,,Metzker,M.,Lee,,Abramzon,S.,Adams,C.,Alder,J.,
Allen,C.,Allen,H.,Alsbrooks,S.,Amlin,A.,Anguiano,D.,
Anylebech,V.,Aoyagi,A.,Ayodeji,M.,Baca,E.,Baden,H.,
Baldwin,D.,Bandaranaike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,
Biswal,K.,Blair,J.,Blankenburg,K.,Blyth,P.,Brown,M.,
Bryant,N.,Burnay,C.,Burch,P.,Burrell,K.,Calderton,E.,
Cardenas,V.,Carter,K.,Cavazos,I.,Ceasar,H.,Center,A.,
Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,
Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Cree,A.,D'Souza,L.,
Davila,M.L.,Davis,C.,Day-Carroll,L.,De Anda,C.,Dederich,D.,
Delgado,O.,Denson,S.,Deramo,C.,Ding,Y.,Dinh,H.,Divya,K.,

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., George, J., E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lourenshaw, L., Louised, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawlin, S., McLeod, M., McNeill, T., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwackelme, O., Okunolu, G., Olarunsgoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C.D., Smales, D., Sneed, A., Sodergren, E., Song, X., Z., Sorrelle, R., Sosa, D., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tinney, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wlasczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, D., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G. and Gibbs, R.A.

Unpublished
Direct Submission
2 (bases 1 to 153671)
Rat Genome Sequencing Consortium.
Submitted (14-AUG-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GJYZ
Center clone name: CH230-10D6
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 72132 bases at least Q40
Consensus quality: 78027 bases at least Q30
Consensus quality: 81644 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 77 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1088: contig of 1088 bp in length
* 1089 1188: gap of unknown length

1189 2277: contig of 1089 bp in length
2278 2377: gap of unknown length
2378 3469: contig of 1092 bp in length
3470 3569: gap of unknown length
3570 4600: contig of 1031 bp in length
4601 4700: gap of unknown length
4701 5725: contig of 1025 bp in length
5726 5825: gap of unknown length
5826 6894: contig of 1069 bp in length
6895 6995: gap of unknown length
6996 8368: contig of 1374 bp in length
8369 9821: gap of unknown length
9822 9922: contig of 1353 bp in length
9923 11138: gap of unknown length
11139 11238: contig of 1217 bp in length
11239 12775: contig of 1537 bp in length
12776 12875: gap of unknown length
12876 13883: contig of 1008 bp in length
13884 13983: gap of unknown length
13984 15301: contig of 1218 bp in length
15302 16748: gap of unknown length
16749 16848: contig of 1447 bp in length
16849 18207: gap of unknown length
18208 18307: contig of 1359 bp in length
18308 20081: gap of unknown length
20082 20181: contig of 1774 bp in length
20182 21605: gap of unknown length
21606 21705: contig of 1424 bp in length
21706 23088: gap of unknown length
23089 23188: contig of 1383 bp in length
23189 24314: gap of unknown length
24315 24414: contig of 1126 bp in length
24415 25833: gap of unknown length
25834 25933: contig of 1419 bp in length
25934 27438: gap of unknown length
27439 27538: contig of 1505 bp in length
27539 28737: gap of unknown length
28738 28837: contig of 1199 bp in length
28838 30058: gap of unknown length
30059 30158: contig of 1221 bp in length
30159 31312: gap of unknown length
31313 31412: contig of 1154 bp in length
31413 32876: gap of unknown length
32877 32976: contig of 1464 bp in length
32977 34532: gap of unknown length
34533 34632: gap of 1556 bp in length
34633 35959: gap of unknown length
35960 37062: contig of 1327 bp in length
37063 37162: gap of unknown length
37163 38287: gap of 1003 bp in length
38287 38387: gap of unknown length
38388 39705: contig of 1125 bp in length
39706 39805: gap of unknown length
39806 40990: contig of 1318 bp in length
40991 41090: gap of unknown length
41091 42586: contig of 1185 bp in length
42587 42686: gap of unknown length
42687 43798: contig of 1456 bp in length
43799 43898: gap of unknown length
43899 45415: contig of 1112 bp in length
45416 45515: gap of unknown length
45516 46835: contig of 1517 bp in length
46836 46935: gap of unknown length
46936 48361: contig of 1320 bp in length
48362 48461: gap of unknown length
48462 50826: contig of 1426 bp in length
50827 50926: gap of unknown length
50927 52239: contig of 2365 bp in length
52239 52339: gap of unknown length
52340 52440: contig of 1313 bp in length
52440 53729: gap of unknown length
53729 53729: contig of 1390 bp in length

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

```

* 53730 53829: gap of unknown length
* 53830 55495: contig of 1666 bp in length
* 53836 55595: gap of unknown length
* 55496 57455: contig of 1860 bp in length
* 57456 57555: gap of unknown length
* 57556 59157: contig of 1602 bp in length
* 59158 59258: gap of unknown length
* 59259 60521: contig of 1264 bp in length
* 60522 60621: gap of unknown length
* 60622 61793: contig of 1172 bp in length
* 61794 61893: gap of unknown length
* 61894 63941: contig of 2048 bp in length
* 63942 64041: gap of unknown length
* 64042 66372: contig of 2331 bp in length
* 66373 66472: gap of unknown length
* 66473 67931: contig of 1459 bp in length
* 67932 68031: gap of unknown length
* 68032 70210: contig of 2179 bp in length
* 70211 70310: gap of unknown length
* 70311 72062: contig of 1752 bp in length
* 72063 72162: gap of unknown length
* 72163 73317: contig of 1155 bp in length
* 73318 74417: gap of unknown length
* 74418 74757: contig of 1340 bp in length
* 74758 74857: gap of unknown length
* 74858 76526: contig of 1669 bp in length
* 76527 76626: gap of unknown length
* 76627 77979: contig of 1353 bp in length
* 77980 78079: gap of unknown length
* 78080 80739: contig of 2660 bp in length
* 80740 80839: gap of unknown length

```

```

Query Match      92.0%; Score 18.4; DB 1; Length 153671;
Best Local Similarity 95.0%; Pred. No. 39;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 TTACTTCTACTTTTACTT 20
Db 108998 TTACTTCTACTTTTACTT 109017

```

```

RESULT 14
AC128449
LOCUS      AC128449      155877 bp      DNA      linear      HTG 19-JUL-2002
DEFINITION Rattus norvegicus clone CH230-117D11, *** SEQUENCING IN PROGRESS

```

```

ACCESSION   AC128449
VERSION     AC128449.1 GI:21909143
KEYWORDS    HTG: HTGS_PHASE1.
SOURCE      Rattus norvegicus.
ORGANISM    Rattus norvegicus.
            Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

```

```

REFERENCE   1 (bases 1 to 155877)
AUTHORS    Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C.,
            Alstbrooks,S.L., Amaralungu,H.C., Are,J.T., Ayele,M., Banks,T.,
            Barbato,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D.,
            Bouck,J., Burich,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
            Buhey,C., Burck,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
            Carro,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
            Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
            Cleveland,C.D., Cox,C., Coyle,M.D., Dahorne,S.R., David,R.,
            Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
            Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinu,H.H.,
            Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
            Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
            Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Franz,P.,
            Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
            Gorisli,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
            Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
            Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B.,
            Homsi,F., Howard,S., Huber,D., Hulyk,S., Hume,J., Jackson,L.E.,

```

```

TITLE      JOURNAL
REFERENCE  2 (bases 1 to 155877)
AUTHORS    Worley,K.C.
TITLE      Direct Submission
JOURNAL    Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA

```

```

COMMENT

```

```

Jacobson,B., Jia,Y., Johnson,R., Jolyet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lien,C., Liu,J., Liu,W., Louised,H.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Meshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwu,S., Ogulu,M., Okwundu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojokan,I., Rolfe,M., Ruiz,S., Saverly,G.,
Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
Sodergren,E., Sonalkhe,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabori,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 155877)
Worley,K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GYWL
Center clone name: CH230-117D11
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 73163 bases at least Q40
Consensus quality: 82138 bases at least Q30
Consensus quality: 87186 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 85 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1062: contig of 1062 bp in length
* 1063 1162: gap of unknown length
* 1163 2171: contig of 1009 bp in length
* 2172 2271: gap of unknown length
* 2272 3321: contig of 1050 bp in length
* 3322 3421: gap of unknown length
* 3422 4463: contig of 1042 bp in length
* 4464 4563: gap of unknown length
* 4564 5587: contig of 1024 bp in length
* 5588 5687: gap of unknown length
* 5688 6713: contig of 1025 bp in length
* 6713 6812: gap of unknown length
* 6812 7822: contig of 1010 bp in length
* 7823 7922: gap of unknown length
* 7923 7953: contig of 1831 bp in length
* 7954 9853: gap of unknown length
* 9854 10887: contig of 1034 bp in length

```

```

* 10888 10987: gap of unknown length
* 10988 12401: contig of 1414 bp in length
* 12402 12501: gap of unknown length
* 12502 14037: contig of 1536 bp in length
* 14038 14137: gap of unknown length
* 14138 15197: contig of 1060 bp in length
* 15198 15297: gap of unknown length
* 15198 16644: contig of 1347 bp in length
* 15298 16744: gap of unknown length
* 16645 18177: contig of 1433 bp in length
* 18178 18277: gap of unknown length
* 18278 19658: contig of 1381 bp in length
* 19659 19758: gap of unknown length
* 19759 20806: contig of 1048 bp in length
* 20807 20906: gap of unknown length
* 20907 22312: contig of 1406 bp in length
* 22313 22412: gap of unknown length
* 22413 23922: contig of 1510 bp in length
* 23923 24022: gap of unknown length
* 24023 25433: contig of 1411 bp in length
* 25434 25533: gap of unknown length
* 25534 26809: contig of 1276 bp in length
* 26810 26910: gap of unknown length
* 26910 28404: contig of 1495 bp in length
* 28405 28504: gap of unknown length
* 28505 29670: contig of 1166 bp in length
* 29671 29770: gap of unknown length
* 29771 30994: contig of 1224 bp in length
* 30995 31094: gap of unknown length
* 31095 32675: contig of 1581 bp in length
* 32676 32776: gap of unknown length
* 32776 34130: contig of 1355 bp in length
* 34131 34230: gap of unknown length
* 34231 35489: contig of 1259 bp in length
* 35490 35589: gap of unknown length
* 35590 37026: contig of 1437 bp in length
* 37027 37126: gap of unknown length
* 37127 38796: contig of 1670 bp in length
* 38797 38896: gap of unknown length
* 38897 40001: contig of 1105 bp in length
* 40002 40101: gap of unknown length
* 40102 41115: contig of 1014 bp in length
* 41116 41215: gap of unknown length
* 41216 42569: contig of 1354 bp in length
* 42570 42669: gap of unknown length
* 42670 43830: contig of 1161 bp in length
* 43831 43930: gap of unknown length
* 43931 45281: contig of 1351 bp in length
* 45282 45381: gap of unknown length
* 45382 47441: contig of 2060 bp in length
* 47442 47541: gap of unknown length
* 47542 49026: contig of 1485 bp in length
* 49027 49126: gap of unknown length
* 49127 50263: contig of 1137 bp in length
* 50264 50363: gap of unknown length
* 50364 52441: contig of 2078 bp in length
* 52442 52541: gap of unknown length
* 52542 53981: contig of 1440 bp in length
* 53982 54081: gap of unknown length
* 54082 55502: contig of 1421 bp in length
* 55503 55602: gap of unknown length
* 55603 56893: contig of 1291 bp in length
* 56894 56993: gap of unknown length
* 56994 58953: contig of 1960 bp in length
* 58954 59053: gap of unknown length
* 59054 60351: contig of 1298 bp in length
* 60352 60451: gap of unknown length
* 60452 61848: contig of 1397 bp in length
* 61849 61948: gap of unknown length
* 61949 64132: contig of 2184 bp in length
* 64133 64232: gap of unknown length
* 64233 65624: contig of 1392 bp in length
* 65625 65724: gap of unknown length

```

```

* 65725 66795: contig of 1071 bp in length
* 66796 66895: gap of unknown length
* 66896 68161: contig of 1266 bp in length
* 68162 68261: gap of unknown length
* 68262 69390: contig of 1129 bp in length
* 69391 69490: gap of unknown length
* 69491 70785: contig of 1295 bp in length
* 70786 70885: gap of unknown length
* 70886 73269: contig of 2384 bp in length
* 73270 73369: gap of unknown length
* 73370 75050: contig of 1681 bp in length
* 75051 75150: gap of unknown length
* 75151 76285: contig of 1135 bp in length
* 76286 76385: gap of unknown length
* 76386 78166: contig of 1781 bp in length
* 78167 78266: gap of unknown length
* 78267 80185: contig of 1919 bp in length
* 80186 80285: gap of unknown length
* 80286 81681: contig of 1396 bp in length
* 81682 81781: gap of unknown length
* 81782 83340: contig of 1559 bp in length
* 83341 83440: gap of unknown length
* 83441 85436: contig of 1996 bp in length
* 85437 85536: gap of unknown length

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 1; Length 155877;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TTAGTCTTCTAGTTTCTAGTT 20
Db 46393 TTAGTCTTCTAGTTCTTCTAGTT 46412

RESULT 15
AC087327
LOCUS AC087327 160264 bp DNA linear HTG 17-JUL-2001
DEFINITION Trypanosoma brucei chromosome IV clone RRC193-3112, *** SEQUENCING
IN PROGRESS ***, 1 ordered piece.
ACCESSION AC087327 6 GI:13378168
VERSION AC087327.6
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Trypanosoma brucei.
ORGANISM Trypanosoma brucei.
REFERENCE 1 (bases 1 to 160264)
AUTHORS El-Sayed,N.M., Ghedin,E., Song,J., Larkin,C., Wanless,D., Jones,K.,
Peterson,J., Hou,L., Zhao,H., Mason,T., Miltischer,J., Pai,G., Van
Aken,S., Uteback,T., Khalak,H.G., Gerard,C., Leech,V., Ullu,E.,
Melville,S., White,O., Adams,M.D., Donelson,J.E. and Fraser,C.M.
Trypanosoma brucei GUTat10.1 RRC193-3112 BAC genomic sequence
JOURNAL Unpublished
TITLE 2 (bases 1 to 160264)
REFERENCE El-Sayed,N.M., Khalak,H. and Adams,M.D.
AUTHORS Direct Submission
JOURNAL Submitted (25-DEC-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Mar 19, 2001 this sequence version replaced gi:13277450.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 160264: contig of 160264 bp in length.
FEATURES
location/Qualifiers
1..160264
/organism="Trypanosoma brucei"
/isolate="GUTat10.1"

```

/db_xref="taxon:5691"
 /chromosome="IV"
 /clone="RPC193-3112"
 BASE COUNT 41037 a 34663 c 39128 g 45436 t
 ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 160264;
 Best Local Similarity 95.0%; Pred. No. 39;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Caps 0;
 QY 1 TTAGTCTTAGTTTAGTT 20
 ||||| |||||
 Db 87884 TTAGTTTAGTTTAGTT 87903

Search completed: April 6, 2003, 11:45:30
 Job time : 428.5 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 10:34:07 ; Search time 243 Seconds
(without alignments)
1069.066 Million cell updates/sec

Title: US-09-980-265-10

Perfect score: 20

Sequence: 1 ttagtcttagtttagt 20

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 1994485 segs, 6494577260 residues

Total number of hits satisfying chosen parameters: 3988970

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl_NoHTG:*
1: gb_ba:*
2: gb_in:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: em_ba:*
15: em_fun:*
16: em_hum:*
17: em_in:*
18: em_mu:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_vl:*
29: em_htg_hum:*
30: em_htg_inv:*
31: em_htg_other:*
32: em_htg_mus:*
33: em_htg_pin:*
34: em_htg_rod:*
35: em_htg_mam:*
36: em_htg_vtl:*
37: em_sy:*
38: em_hngo_hum:*
39: em_hngo_mus:*
40: em_hngo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	20	5	AX057376	AX057376 Sequence
2	19	95.0	123078	7	AP004902	AP004902 Lotus jap
3	18.4	92.0	20	5	AX057368	AX057368 Sequence
4	18.4	92.0	20	5	AX057372	AX057372 Sequence
5	18.4	92.0	20	5	AX057375	AX057375 Sequence
6	18.4	92.0	20	5	AX057377	AX057377 Sequence
7	18.4	92.0	20	5	AX057379	AX057379 Sequence
8	18.4	92.0	20	5	AX057380	AX057380 Sequence
9	18.4	92.0	65	5	AX485115	AX485115 Sequence
10	18.4	92.0	2280	1	AF306797	AF306797 Vibrio ch
11	18.4	92.0	484	7	AF170083	AF170083 Candida a
12	18.4	92.0	5453	5	AX252115	AX252115 Sequence
13	18.4	92.0	5453	5	AX344477	AX344477 Sequence
14	18.4	92.0	5453	5	AX348882	AX348882 Sequence
15	18.4	92.0	7238	5	AX345350	AX345350 Sequence
16	18.4	92.0	7353	5	AX344974	AX344974 Sequence
17	18.4	92.0	7353	5	AX348495	AX348495 Sequence
18	18.4	92.0	10682	5	AX344677	AX344677 Sequence
19	18.4	92.0	11670	5	AX281283	AX281283 Sequence
20	18.4	92.0	16258	5	AX348425	AX348425 Sequence
21	18.4	92.0	16258	5	AX348808	AX348808 Sequence
22	18.4	92.0	18817	5	AX281384	AX281384 Sequence
23	18.4	92.0	18817	5	AX348594	AX348594 Sequence
24	18.4	92.0	21537	5	AX346900	AX346900 Sequence
25	18.4	92.0	28928	2	CEZC410	268270 Caenorhabdi
26	18.4	92.0	97191	7	AP004499	AP004499 Lotus jap
27	18.4	92.0	106949	2	CEV52B11A	AL037654 Caenorhab
28	18.4	92.0	112063	8	AC091485	AC091485 Homo sapi
29	18.4	92.0	115046	7	AP004907	AP004907 Lotus jap
30	18.4	92.0	118873	8	AL157878	AL157878 Human DNA
31	18.4	92.0	121280	8	AC087045	AC087045 Homo sapi
32	18.4	92.0	131741	7	AP004916	AP004916 Lotus jap
33	18.4	92.0	137908	8	AL607024	AL607024 Mouse DNA
34	18.4	92.0	154259	8	AC013284	AC013284 Homo sapi
35	18.4	92.0	170623	8	AL391375	AL391375 Human DNA
36	18.4	92.0	173456	8	AL450334	AL450334 Human DNA
37	18.4	92.0	173456	8	AC055758	AC055758 Homo sapi
38	18.4	92.0	175867	2	AC007976	AC007976 Drosophila
39	18.4	92.0	184621	2	AC007440	AC007440 Drosophila
40	18.4	92.0	185096	8	AL355140	AL355140 Human DNA
41	18.4	92.0	186199	8	AC093911	AC093911 Homo sapi
42	18.4	92.0	186308	8	AC026739	AC026739 Homo sapi
43	18.4	92.0	189944	8	AC009563	AC009563 Homo sapi
44	18.4	92.0	196912	9	AL513346	AL513346 Mouse DNA
45	18.4	92.0	214668	8	AC073415	AC073415 Homo sapi

ALIGNMENTS

RESULT 1
AX057376
LOCUS AX057376 20 bp DNA
DEFINITION Sequence 10 from Patent WO0075304.
ACCESSION AX057376
VERSION AX057376.1 GI:12310117
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
artificial sequences.
REFERENCE
1 (bases 1 to 20)
AUTHORS Bachy,M., Sodyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 10 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
Location/Qualifiers

source 1. .20
/organism="synthetic construct"
/db.xref="taxon:32630"
/note="Oligonucleotide"

BASE COUNT 3 a 1 c 3 g 13 t

ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTTTACTT 20
|||||

Db 1 TTAGTCTTACTTTTACTT 20

RESULT 2
AP004902/c 123078 bp DNA linear PLN 19-JUL-2002
LOCUS Lotus japonicus genomic DNA, chromosome 2, clone:LTJ04G24, TM0060,
DEFINITION complete sequence.
ACCESSION AP004902
VERSION AP004902.1 GI:21907918
KEYWORDS HTG.
SOURCE Lotus japonicus DNA, clone_lib:LJT library clone:LTJ04G24.
ORGANISM Lotus japonicus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteeae;
Lotus.

REFERENCE 1
AUTHORS Kaneko,T., Nakamura,Y., Asamizu,E., Kato,T., Sato,S. and Tabata,S.
TITLE Structural Analysis of a Lotus japonicus Genome. I. Sequence
Features and Mapping of Sixty-six TAC clones which cover the 6.7 Mb
Regions of the Genome
Unpublished
REFERENCE 2 (bases 1 to 123078)
AUTHORS Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (26-MAR-2002) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research, 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:yn@kazusa.or.jp,
url:http://www.kazusa.or.jp, Tel:81-438-52-3935,
Fax:81-438-52-3934)

FEATURES
source Location/Qualifiers
1. .123078
/organism="Lotus japonicus"
/db.xref="taxon:34305"
/chromosome="2"
/clone="LJT04G24"
/clone_lib="LJT library"
/note="TAC clone:TM0060"

BASE COUNT 40721 a 22328 c 21292 g 38737 t

ORIGIN

Query Match 95.0%; Score 19; DB 7; Length 123078;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TAGTCTTACTTTTACTT 20
|||||

Db 77977 TAGTCTTACTTTTACTT 77959

RESULT 3
AX057368 20 bp DNA linear PAT 17-JAN-2001
LOCUS AX057368
DEFINITION Sequence 2 from Patent WO0075304.
ACCESSION AX057368
VERSION AX057368.1 GI:12310109
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct

artificial sequences.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Trannoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 2 14-DEC-2000;
Aventis Pasteur (FR)

FEATURES
source Location/Qualifiers
1. .20
/organism="synthetic construct"
/db.xref="taxon:32630"
/note="Oligonucleotide"

BASE COUNT 4 a 0 c 3 g 13 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTTTACTT 20
|||||

Db 1 TTAGTCTTACTTTTACTT 20

RESULT 4
AX057372 20 bp DNA linear PAT 17-JAN-2001
LOCUS AX057372
DEFINITION Sequence 6 from Patent WO0075304.
ACCESSION AX057372
VERSION AX057372.1 GI:12310113
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Trannoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 6 14-DEC-2000;
Aventis Pasteur (FR)

FEATURES
source Location/Qualifiers
1. .20
/organism="synthetic construct"
/db.xref="taxon:32630"
/note="Oligonucleotide"

BASE COUNT 3 a 0 c 3 g 14 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTTTACTT 20
|||||

Db 1 TTAGTCTTACTTTTACTT 20

RESULT 5
AX057375 20 bp DNA linear PAT 17-JAN-2001
LOCUS AX057375
DEFINITION Sequence 9 from Patent WO0075304.
ACCESSION AX057375
VERSION AX057375.1 GI:12310116
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct

REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Trannoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 9 14-DEC-2000;
Aventis Pasteur (FR)

FEATURES
source Location/Qualifiers
1. .20
/organism="synthetic construct"

BASE COUNT 4 a 1 c 3 g 12 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTGTAGTT 20
Db 1 TTAGTCTTAGTTTGTAGTT 20

RESULT 6
AX057377 20 bp DNA linear PAT 17-JAN-2001
LOCUS AX057377
DEFINITION Sequence 11 from Patent WO0075304.
ACCESSION AX057377
VERSION AX057377.1 GI:12310118
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy, M., Sodoyer, R. and Tranoy, E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 11 14-DEC-2000;
Aventis Pasteur (FR)

FEATURES
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide"

BASE COUNT 3 a 2 c 3 g 12 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTGTAGTT 20
Db 1 TTAGTCTTAGTTTGTAGTT 20

RESULT 7
AX057379 20 bp DNA linear PAT 17-JAN-2001
LOCUS AX057379
DEFINITION Sequence 13 from Patent WO0075304.
ACCESSION AX057379
VERSION AX057379.1 GI:12310120
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy, M., Sodoyer, R. and Tranoy, E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 13 14-DEC-2000;
Aventis Pasteur (FR)

FEATURES
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide"

BASE COUNT 3 a 1 c 4 g 12 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTGTAGTT 20
Db 1 TTAGTCTTAGTTTGTAGTT 20

RESULT 8
AX057380 20 bp DNA linear PAT 17-JAN-2001
LOCUS AX057380
DEFINITION Sequence 14 from Patent WO0075304.
ACCESSION AX057380
VERSION AX057380.1 GI:12310121
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy, M., Sodoyer, R. and Tranoy, E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 14 14-DEC-2000;
Aventis Pasteur (FR)

FEATURES
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide"

BASE COUNT 3 a 0 c 4 g 13 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTGTAGTT 20
Db 1 TTAGTCTTAGTTTGTAGTT 20

RESULT 9
AX485115/6 65 bp DNA linear PAT 16-AUG-2002
LOCUS AX485115
DEFINITION Sequence 2415 from Patent WO02053728.
ACCESSION AX485115
VERSION AX485115.1 GI:22319399
KEYWORDS
SOURCE Candida albicans.
ORGANISM Candida albicans.
REFERENCE 1
AUTHORS Roemer, T., Jiang, B., Boone, C., Bussey, H. and Ohlsen, K. L.
TITLE Gene disruption methodologies for drug target discovery
JOURNAL Patent: WO 02053728-A 2415 11-JUL-2002;
Eli Lilly Pharmaceuticals, Inc. (US)

FEATURES
source 1..65
/organism="Candida albicans"
/db_xref="taxon:5476"

BASE COUNT 37 a 6 c 16 g 6 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 65;
Best Local Similarity 95.0%; Pred. No. 9.5e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTGTAGTT 20
Db 46 TTAGTCTTAGTTTGTAGTT 27

RESULT 10
AF306797/c 2280 bp DNA linear BCT 23-JUL-2001
LOCUS AF306797
DEFINITION Vibrio cholerae strain SCE256 toxin co-regulated pilus biosynthesis

protein F (tcpF) and toxin co-regulated pilus virulence regulatory protein (tcpN) genes, complete cds.

AF306797
AF306797.1 GI:11139681

VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
SOURCE

Vibrio cholerae.
Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
1 (bases 1 to 2280)
Mukhopadhyay, A.K., Chakraborty, S., Takeda, Y., Nair, G.B. and Berg, D.E.
Characterization of VPI pathogenicity island and CTXphi prophage in environmental strains of Vibrio cholerae
J. Bacteriol. 183 (16), 4737-4746 (2001)
21359315
11466276
2 (bases 1 to 2280)
Mukhopadhyay, A.K., Chakraborty, S., Shinnada, T., Takeda, Y., Nair, G.B. and Berg, D.E.
Direct Submission
Submitted (20-SEP-2000) Molecular Microbiology, Washington University Medical School, 660 South Euclid Ave., Saint Louis, MO 63110, USA

Location/Qualifiers
1..2280
/organism="Vibrio cholerae"
/strain="SCE256"
/db_xref="taxon:666"
189..1199
/gene="tcpF"
189..1199
/gene="tcpF"
/note="tcpF"
/note="type III"
/codon_start=1
/transl_table=11
/product="toxin co-regulated pilus biosynthesis protein F"
/protein_id="AA331794.1"
/db_xref="GI:11139682"
/translation="MKLKLSSSIILLNSALAFNDNYSSTSTVYSSSTONGDOAGS
SGNRYRPFACLSLGPSSLSYCTVAFPPDSKEFYGVPSIEGVTRPDPFNK
LMTGAYTELESKYLKPYLGQOTPDMAIPPDQOTLVNLSKNGDLKESKDYNSGRF
LTKTDGEVYIAIGNVAEDSPKLPDIYKKNLSDELPSNAVCKLTKTKTY
VTQDYSRRDPAFTLIPDEKRENDIPTRVYEVLYKRGSGDGIIRNPKTTPTIYI
DPRSPNSTRFKSDSLHNGDSSDNKTIQLYTPDSKYSTLKALPLEFDCRSQCEGV
SD"
1409..2242
/gene="tcpN"
1409..2242
/gene="tcpN"
/note="tcpN"
/note="toxin"
/codon_start=1
/transl_table=11
/product="toxin co-regulated pilus virulence regulatory protein"
/protein_id="AA331795.1"
/db_xref="GI:11139683"
/translation="MRENNTSFLVNYVDLKKPDTYAFNKVFIIDYKIFWINKSARKY
DKNCLSYTTASSVVLKKNLSIORSIMSLSDESISVCVLTINKFVNLRHLYGQD
LMTIRNLNEKKDILLNMCLENDISVGEIYSTVDQOYSEDPLKIFPSGFSVEYKY
NSIEFTDDLDAMEKISCIYKSDITRRMRADICGELRTNMLLKKLESGVAFRELI
NSIRISYSISLAKTGFKIKQIAYSGFASVSFTVFSTMNVAPESTLFLMLTGAVE
E"

BASE COUNT 773 a 334 c 395 g 778 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 2280;
Best Local Similarity 95.0%; Pred. No. 3.3e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTAGTTTGTAGT 20
||||| |||||||||
DB 1308 TTAGTTTGTAGTTTGTAGT 1289

RESULT 11
AF170083/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REMARK
COMMENT
FEATURES
SOURCE

Candida albicans putative transcriptional repressor (SSN6) gene, complete cds.
AF170083
AF170083.2 GI:18767667

Candida albicans.
Candida albicans
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
1 (bases 1 to 4844)
Hwang, C.-S. and Kang, S.-O.
Molecular cloning of the gene encoding Candida albicans Prf1 homologous to Saccharomyces cerevisiae Ssn6
Unpublished
2 (bases 1 to 4844)
Hwang, C.-S. and Kang, S.-O.
Direct Submission
Submitted (19-JUL-1999) Department of Microbiology, College of Natural Science, Seoul National University, Shinlim-Dong, Kwanak-Ku, Seoul 151-742, South Korea
3 (bases 1 to 4844)
Kang, S.-O.
Direct Submission
Submitted (20-FEB-2002) School of Biological Sciences, Institute of Microbiology, Seoul National University, Shinlim-Dong, Kwanak-Ku, Seoul 151-742, South Korea
Sequence update by submitter
On Feb 20, 2002 this sequence version replaced gi:18000300.

Location/Qualifiers
1..4844
/organism="Candida albicans"
/strain="ATCC10231"
/db_xref="ATCC:10231"
/db_xref="taxon:5476"
1132..24389
/gene="SSN6"
1132..24389
/gene="SSN6"
/product="putative transcriptional repressor"
1132..4389
/gene="SSN6"
/note="Prf1p; similar to Saccharomyces cerevisiae Ssn6"
/codon_start=1
/transl_table=12
/product="putative transcriptional repressor"
/protein_id="AAL54912.2"
/db_xref="GI:18767668"
/translation="MYTAHTTIKQOQOQHPPPLNGGLHAGSAPPNSHEAAIAQOQ
OOOHHNGMGITVAAAASANOAAVOARAQOQOQOQOOLPSSAALNETVSTWLA
IGSLAESIGDIERATASYNALPHSNPQDILVTKINTYRSKQOPLKAELVQDALNF
HYNGETWGLGHCYTLMDLDQARAAYORALFYLENPNVPRKIMHIGITLXDYGSLE
YAEAFVRVLDLPNDFKANETIFRLGITYKHOGKLPALDECQYILNPPHPLTQPD
VWFQISGYEEOQDNMGADAYEKVLIQHPHAKVLAQGLCLTQAESNSPPANPANA
OPKPPQOQDTLLKYLKQSLVEDQSDAHSMWYLGHEVIMRGDTAAYAFQOAVRD
ARNPTWMSIGVLYYQISOYRQADQVYRAIRLNPVISEWVWYGLTYAECCNNOISDA
LADLAVYRAERLDPDNNPHKARLDQTLTYRQOQGNHPPQPPSSQOPLPQGMVLES
TQOQOQOQOQPPPPPPPOOQLOHOSOSOPQOQPPQOQPSLOHOSLPPQOI
OPHQQAAPFLVNOQOSPPPHLMNLGOPEQOQROLPHLPHTQOQSOIQEKPPQOE
OPHHQPPPPPOHQQOSQOPPPHPOHONOSPOLAOLPHHSNPPANPHGAPQOQT
GLPDLHNSANIILIASVOYPOQOQPHIAPVROGVNHPVSIASAPRPETTLPO
INNPNESTTQVOPOLKKEPKRPATVSAPVPAIKVQDOVTTQIESAPAAAVASAPASA
PVGDITDVTSTTTPATSTTADAVPVSYSGVAGPAPVNOEKYKPRDEQIVSOYEKVE
SQEPETPAPPAALATAPLEPAITDKVDVMAVSKSATPLPOSIDQNTRVSBATAP
ESGKHLDLDEKNDKILKRPVYETTESVPVNOPEVKEKNEKEVEPPPOPSSEKRE
KEYNGSIRKPLENESKVDIPQSSNTTQANEAKSGSEETRKDTYTKPAKQGVKCVI
PSETETVSKPDVENKDKKEDKDEVDNEDDVKKDNDENPPARKTEEDENYDDE"

BASE COUNT 1722 a 1040 c 736 g 1346 t

ORIGIN

Query Match	92.0%;	Score 18.4;	DB 7;	Length 484;
Best Local Similarity	95.0%;	Pred. No. 2.6e02;		
Matches 19;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	TTAGTCTTACTTTTACTT	20	
Db	392	TTAGTCTTACTTTTACTT	373	

	RESULT 12			
LOCUS	AX252115			
DEFINITION	AX252115	5453 bp	DNA	linear
ACCESSION	Sequence 376 from Patent WO0168911.			PAT 05-OCT-2001
VERSION	AX252115			
KEYWORDS	AX252115.1 GI:15985470			
SOURCE	.			
ORGANISM	synthetic construct. synthetic construct. artificial sequences.			
REFERENCE	1 (bases 1 to 5453)			
AUTHORS	Olek A., Piepenbrock,C. and Berlin,K.			
TITLE	Diagnosis of diseases associated with the cell cycle			
JOURNAL	Patent: WO 0168911-A 376 20-SRP-2001; Epigenomics AG (DE) Location/Qualifiers 1..5453 /organism="synthetic construct" /db_xref="taxon:32630" /note="Chemically treated genomic DNA (Homo sapiens)"			
FEATURES				
source				
BASE COUNT	1494 a 101 c 1213 g 2645 t			
ORIGIN				

Query Match	92.0%	Score 18.4	DB 5	Length 5453
Best Local Similarity	95.0%	Pred. No. 2.5e+02		
Matches 19	Conservative	0	Mismatches 1	Indels 0
QY	1	TTATGTCCTAGCTTTTACTT	20	
db	983	TTAGCTTTTAGCTTTTACTT	1002	

```

RESULT 13
AX344477
LOCUS AX344477 5453 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 324 from Patent WO0200926.
ACCESSION AX344477
VERSION AX344477.1 GI:18492365
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with signal transduction
JOURNAL Patent: WO 0200926-A 324 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
source
location/Qualifiers
1..5453
/organism="Synthetic Construct"
/db_xref="taxon:37630"
/note="Chemically treated genomic DNA (Homo sapiens)"
BASE COUNT 1494 a 101 c 1213 g 2645 t
ORIGIN

```

	Query Match	Score	DB	Length
	92.0%	18.4	5	5453
	Best Local Similarity	95.0%		
	Pred. No. 2.5e+02			
Matches	19; Conservative	0; Mismatches	1; Indels	0; Gaps
Qy	1	TTAGTCTTAGCTTTTACGT	20	
Db	983	TTAGTTTACGTTTACGTT	1002	

OY	1	TTAGTCTTAGTTTTACTT	20
db	983	TTAGTTTTAGTTTTACTT	1002
LOCUS	AX348882	5453 bp	DNA
DEFINITION	Sequence 340 from Patent WO0202807.		linear
ACCESSION	AX348882		PAT 06-FEB-2002
VERSION	AX348882.1	GI:18614917	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE	1		
AUTHORS	Olek,A., Piepenbrock,C. and Berlin,K.		
TITLE	Diagnosis of diseases associated with cell signalling		
JOURNAL	Patent: WO 0202807-A 340 10-JAN-2002;		
FEATURES	Epigenomics AG (DE)		
source	location/Qualifiers		
	1..5453		
	/organism="synthetic construct"		
	/db_xref="taxon:32630"		
	/note="Chemically treated genomic DNA (Homo sapiens)"		
BASE COUNT	1494 a	101 c	1213 g
ORIGIN		2645 t	
Query Match	92.0%:	Score 18.4:	DB 5:
Best Local Similarity	95.0%:	Pred. No. 2.5e+02:	Length 5453:
Matches 19:	Conservative 0:	Mismatches 1:	Indels 0:
		Gaps 0:	

db	QY	1	TTAGTCTTACTTTTACTT	20	92.0%: Score 18.4; DB 5; Length 7238;
6279	TTAGTCTTACTTTTACTT	6298			Best Local Similarity 95.0%; Pred. No. 2.3e+02;
	Matches 19;	Conservative	0;	Mismatches	1; Indels 0; Gaps 0;

LOCUS	AX345350	7238 bp	DNA	linear	PAT 01-FEB-2002
DEFINITION	Sequence 421 from Patent WO0200928.				
ACCESSION	AX345350				
VERSION	AX345350.1	GI:18493236			
KEYWORDS	.				
SOURCE	synthetic construct.				
ORGANISM	synthetic construct				
REFERENCE	1				
AUTHORS	Olek, A., Piepenbrock, C. and Berlin, K.				
TITLE	Diagnosis of diseases associated with the immune system				
JOURNAL	Patent: WO 0200928-A 421 03-JAN-2002;				
FEATURES	Epigenomics AG (DE)				
source	Location/Qualifiers				
	1..7238				
	/organism="synthetic construct"				
	/db_xref="taxon:32630"				
	/note="Chemically treated genomic DNA (Homo sapiens)"				
BASE COUNT	2414 a	84 c	1401 g	3339 t	
ORIGIN					

```
Search completed: April 6, 2003, 11:53:46
Job time : 268 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:21:32 ; Search time 151 Seconds
(without alignments)
298.278 Million cell updates/sec

Title: US-09-980-265-10

Perfect score: 20

Sequence: 1 ttacttcttagtttagtt 20

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N.Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	22	AAF28881 Immunostimulatory
2	19	95.0	20	22	AAF28871 Genetic Immunostim
3	18.4	92.0	20	22	AAF28873 Immunostimulatory
4	18.4	92.0	20	22	AAF28877 Immunostimulatory
5	18.4	92.0	20	22	AAF28880 Immunostimulatory
6	18.4	92.0	20	22	AAF28882 Immunostimulatory
7	18.4	92.0	20	22	AAF28883 Immunostimulatory
8	18.4	92.0	20	22	AAF28885 Immunostimulatory
9	18.4	92.0	5453	24	ABL70450 Chemically treated

10	18.4	92.0	5453	24	AAS61413 Human gene regulat
11	18.4	92.0	5453	24	ABK31481 Signal transductio
12	18.4	92.0	7238	24	ABL32448 Human immune syste
13	18.4	92.0	7353	24	ABL32072 Human immune syste
14	18.4	92.0	7353	24	AAD28362 Human chemically t
15	18.4	92.0	10682	24	ABN80085 Human chemically m
16	18.4	92.0	11670	24	ABL54325 Chemically treated
17	18.4	92.0	16258	24	ABL70376 Chemically treated
18	18.4	92.0	16258	24	ABK40038 Human chemically p
19	18.4	92.0	18817	24	ABL70162 Chemically treated
20	18.4	92.0	18817	24	ABL34495 Human metastasis a
21	18.4	92.0	21537	24	ABL33998 Human immune syste
22	17.4	87.0	711	24	ABO43350 Oligonucleotide fo
23	17.4	87.0	711	24	ABO43351 Oligonucleotide fo
24	17.4	87.0	885	24	ABO49064 Oligonucleotide fo
25	17.4	87.0	885	24	ABO49065 Oligonucleotide fo
26	17.4	87.0	3745	23	ABL26212 Drosophila melanog
27	17.4	87.0	3775	23	ABL26188 Human immune syste
28	17.4	87.0	5198	24	ABL32279 Human immune syste
29	17.4	87.0	6012	24	ABL70464 Chemically treated
30	17.4	87.0	6012	24	ABK31491 Signal transductio
31	17.4	87.0	7500	24	ABL33115 Human immune syste
32	17.4	87.0	8467	24	ABL32108 Human immune syste
33	17.4	87.0	15698	24	ABL34141 Human immune syste
34	17.4	87.0	21667	23	ABL03568 Drosophila melanog
35	16.8	84.0	20	22	AAF28872 Immunostimulatory
36	16.8	84.0	20	22	AAF28874 Immunostimulatory
37	16.8	84.0	20	22	AAF28875 Immunostimulatory
38	16.8	84.0	20	22	AAF28876 Immunostimulatory
39	16.8	84.0	20	22	AAF28878 Immunostimulatory
40	16.8	84.0	20	22	AAF28879 Immunostimulatory
41	16.8	84.0	20	22	AAF28884 Immunostimulatory
42	16.8	84.0	20	22	AAF28886 Immunostimulatory
43	16.8	84.0	20	22	AAF28887 Immunostimulatory
44	16.8	84.0	488	24	ABL93059 Rat metastatic tum
45	16.8	84.0	569	24	ABO48528 Oligonucleotide fo

ALIGNMENTS

RESULT 1	
AAF28881	AAF28881 standard; DNA; 20 BP.
ID	
AC	AAF28881;
XX	
DT	09-MAY-2001 (first entry)
XX	
DE	Immunostimulatory oligonucleotide #10 as vaccine adjuvant.
XX	
KW	Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW	propylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW	cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW	phosphorothioate; ss.
XX	
OS	Synthetic.
XX	
FH	Key
FT	modified_base 1..20
FT	Location/Qualifiers
FT	/tag= a
FT	/note= "contain phosphorothioate internucleotide linkages"
XX	
PN	WO200075304-A1.
XX	
PD	14-DEC-2000.
XX	
PF	08-JUN-2000; 2000WO-FR01566.
XX	
PR	08-JUN-1999; 99FR-0007457.
PR	06-AUG-1999; 99FR-0010378.
XX	

PA (AVET) AVENTIS PASTEUR.
XX
PI Bachy M, Sodayer R, Tranoy E;
XX
XX WPI; 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unethylyated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
XX
SQ Sequence 20 BP; 3 A; 1 C; 3 G; 13 T; 0 other;
XX
Query Match 100.0%; Score 20; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTTGT 20
1 TTAGTCTTAGTTTGT 20
Db 1 TTAGTCTTAGTTTGT 20
XX
RESULT 2
AAF28871
ID AAF28871 standard; DNA: 20 BP.
XX
AC AAF28871;
XX
DT 09-MAY-2001 (first entry)
XX
DE Generic immunostimulatory oligonucleotide as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor; ss.
XX
OS Synthetic.
XX
PN WO200075304-A1.
XX
PD 14-DEC-2000.
XX
PE 08-JUN-2000; 2000WO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVET) AVENTIS PASTEUR.
XX
PI Bachy M, Sodayer R, Tranoy E;
XX
DR WPI; 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Claim 8; Page 17; 30pp; French.
XX
CC This sequence represents a generic example of an immunostimulatory
CC oligonucleotide of the invention which contains at least one sequence

CC 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The oligonucleotides
CC do not contain any CG dinucleotides in which C are unethylyated.
CC The oligonucleotides are used as human immunostimulants and as adjuvants
CC in therapeutic and prophylactic vaccines for human use. They induce
CC proliferation of human lymphocytes, induce secretion of cytokines,
CC especially interleukin-10 or interferon-gamma and increase expression
CC of the CD86 activation marker or the CD25 cytokine receptor on human B
CC lymphocytes.
XX
SQ Sequence 20 BP; 3 A; 1 C; 3 G; 12 T; 1 other;
XX
Query Match 95.0%; Score 19; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 82;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTTGT 20
1 TTAGTCTTAGTTTGT 20
Db 1 TTAGTCTTAGTTTGT 20
XX
RESULT 3
AAF28873
ID AAF28873 standard; DNA: 20 BP.
XX
AC AAF28873;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #2 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
PN WO200075304-A1.
XX
PD 14-DEC-2000.
XX
PE 08-JUN-2000; 2000WO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVET) AVENTIS PASTEUR.
XX
PI Bachy M, Sodayer R, Tranoy E;
XX
DR WPI; 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unethylyated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor

```
CC on human B lymphocytes.
XX
SQ Sequence 20 BP; 4 A; 0 C; 3 G; 13 T; 0 other;
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTACTTTTACTT 20
   ||||| ||||| ||||| |||||
Db 1 TTAGTATTACTTTTACTT 20

RESULT 4
AAF28877
ID AAF28877 standard; DNA; 20 BP.
XX
AC AAF28877;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #6 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /*tag= a
FT /note= "contain phosphorothioate internucleotide
FT linkages"
XX
XX WO200075304-A1.
XX
XX 14-DEC-2000.
XX
XX 08-JUN-2000; 2000WO-FR01566.
XX
XX 08-JUN-1999; 99FR-0007457.
XX
XX 06-AUG-1999; 99FR-0010378.
XX
XX (AVET ) AVENTIS PASTEUR.
XX
XX Bachy M, Sodoyer R, Tranoy E;
XX
XX WPI; 2001-041317/05.
XX
XX
XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
XX vaccines for human use, induce lymphocyte proliferation and cytokine
XX secretion -
XX
XX Example 4; Page 13; 30pp; French.
XX
XX Oligonucleotides AAF28872-AAF28887 represent specific examples of an
XX immunostimulatory oligonucleotide of the invention which contains at
XX least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
XX oligonucleotides do not contain any CG dinucleotides in which C are
XX unmethylated. The oligonucleotides are used as human immunostimulants
XX and as adjuvants in therapeutic and prophylactic vaccines for human use.
XX They induce proliferation of human lymphocytes, induce secretion of
XX cytokines, especially interleukin-10 or interferon-gamma and increase
XX expression of the CD86 activation marker or the CD25 cytokine receptor
XX on human B lymphocytes.
XX
SQ Sequence 20 BP; 3 A; 0 C; 3 G; 14 T; 0 other;
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 TTAGTCTTACTTTTACTT 20
   ||||| ||||| ||||| |||||
Db 1 TTAGTCTTACTTTTACTT 20

RESULT 5
AAF28880
ID AAF28880 standard; DNA; 20 BP.
XX
AC AAF28880;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #9 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /*tag= a
FT /note= "contain phosphorothioate internucleotide
FT linkages"
XX
XX WO200075304-A1.
XX
XX 14-DEC-2000.
XX
XX 08-JUN-2000; 2000WO-FR01566.
XX
XX 08-JUN-1999; 99FR-0007457.
XX
XX 06-AUG-1999; 99FR-0010378.
XX
XX (AVET ) AVENTIS PASTEUR.
XX
XX Bachy M, Sodoyer R, Tranoy E;
XX
XX WPI; 2001-041317/05.
XX
XX
XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
XX vaccines for human use, induce lymphocyte proliferation and cytokine
XX secretion -
XX
XX Example 4; Page 13; 30pp; French.
XX
XX Oligonucleotides AAF28872-AAF28887 represent specific examples of an
XX immunostimulatory oligonucleotide of the invention which contains at
XX least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
XX oligonucleotides do not contain any CG dinucleotides in which C are
XX unmethylated. The oligonucleotides are used as human immunostimulants
XX and as adjuvants in therapeutic and prophylactic vaccines for human use.
XX They induce proliferation of human lymphocytes, induce secretion of
XX cytokines, especially interleukin-10 or interferon-gamma and increase
XX expression of the CD86 activation marker or the CD25 cytokine receptor
XX on human B lymphocytes.
XX
SQ Sequence 20 BP; 4 A; 1 C; 3 G; 12 T; 0 other;
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTACTTTTACTT 20
   ||||| ||||| ||||| |||||
Db 1 TTAGTCTTACTTTTACTT 20

RESULT 6
```

```
AAF28882
ID AAF28882 standard; DNA: 20 BP.
XX
AC AAF28882;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #11 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /tag= a
FT /note= "contain phosphorothioate internucleotide
FT linkages"
XX
PN WO200075304-A1.
XX
PD 14-DEC-2000.
XX
PF 08-JUN-2000; 2000WO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVET ) AVENTIS PASTEUR.
XX
PI Bachy M, Sodoyer R, Tranney E;
XX
DR WPI; 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
XX
SQ Sequence 20 BP; 3 A; 2 C; 3 G; 12 T; 0 other;
XX
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTTCTTAGTT 20
DB 1 TTAGTCTTAGTTCTTAGTT 20

RESULT 7
AAF28883
ID AAF28883 standard; DNA: 20 BP.
XX
AC AAF28883;
XX
DT 09-MAY-2001 (first entry)
XX
```

```
DE Immunostimulatory oligonucleotide #12 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /tag= a
FT /note= "contain phosphorothioate internucleotide
FT linkages"
XX
PN WO200075304-A1.
XX
PD 14-DEC-2000.
XX
PF 08-JUN-2000; 2000WO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVET ) AVENTIS PASTEUR.
XX
PI Bachy M, Sodoyer R, Tranney E;
XX
DR WPI; 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
XX
SQ Sequence 20 BP; 3 A; 1 C; 4 G; 12 T; 0 other;
XX
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTTCTTAGTT 20
DB 1 TTAGTCTTAGTTGTTTAGTT 20

RESULT 8
AAF28885
ID AAF28885 standard; DNA: 20 BP.
XX
AC AAF28885;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #14 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
```


PA (EPIC-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-017470/02.
XX
PT New nucleic acid sequences from chemically modified genes associated
PT with gene regulation, useful for analysing cytosine methylations for
PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
XX
XX
PS Disclosure; SEQ ID No 376; 26pp; English.
XX
CC The invention relates to 224 nucleic acid sequences comprising at least
CC 18 bases of a chemically pretreated gene associated with gene regulation
CC selected from 43 known genes (or complementary sequences). The
CC chemical pretreatment converts cytosine bases unmethylated at the
CC 5-position to uracil or another base with hybridisation behaviour
CC dissimilar to cytosine, to enable analysis of cytosine methylations.
CC The DNA sequences, oligomers (or sets/arrays) and method are
CC useful in the diagnosis of diseases (or predisposition to diseases)
CC associated with gene regulation and in therapy of such diseases, by
CC enabling analysis of the cytosine methylation patterns of such genes,
CC kits are provided. They are especially useful in diagnosis
CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
CC pre-eclampsia, graft versus-host disease. The present sequence is a
CC sequence included in the sequence data for this specification and is
CC associated with the human gene regulation-associated genes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 5453 BP; 1494 A; 101 C; 1213 G; 2645 T; 0 other;
XX
Query Match 92.0%; Score 18.4; DB 24; Length 5453;
Best Local Similarity 95.0%; Pred. NO. 1.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTTGTGTT 20
Db 983 TTAGTTTGTGTTTGTGTT 1002
XX
RESULT 11
ABK31481
ID ABK31481 standard; DNA: 5453 BP.
XX
AC ABK31481;
XX
DT 23-APR-2002 (first entry)
XX
DE Signal transduction associated gene modified complementary DNA #162.
XX
KW Human: signal transduction associated gene; cytosine methylation state;
KW CpG island; signal transduction associated disease; solid tumour; cancer;
KW antitumour; cytostatic; mutant; ds.
XX
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200200926-A2.
XX
PD 03-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP07472.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIC-) EPIGENOMICS AG.

XX
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-147896/19.
XX
PT Oligonucleotide for diagnosis and therapy of diseases associated with
PT signal transduction e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with signal transduction
XX
XX
PS Claim 1; SEQ ID No 324; 24pp; English.
XX
CC The present invention relates to chemically modified DNA sequences of
CC signal transduction associated genes. The DNA sequences are chemically
CC modified using a solution of bisulphite, hydrogen sulphite or
CC disulphite. Also disclosed are oligonucleotides and/or PNA oligomers
CC for detecting the cytosine methylation state (CpG islands) of these
CC genes, and a method for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with signal transduction.
CC The genomic DNA can be obtained from cells or cellular components which
CC contain DNA, e.g. cell lines, biopsies, blood, sputum, stool, urine,
CC cerebral-spinal fluid, tissue embedded in paraffin such as tissue from
CC eyes, intestine, kidney, brain, heart, prostate, lung, breast or liver,
CC histologic object slides, and all their possible combinations. The
CC sequences of the invention are useful for the diagnosis and therapy of
CC diseases associated with signal transduction e.g. solid tumours and
CC cancer. ABK31158-ABK31545 represent chemically pretreated genomic DNA
CC sequences of different genes associated with signal transduction, or
CC their complementary sequences.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from the
CC European Patent Office.
XX
SQ Sequence 5453 BP; 1494 A; 101 C; 1213 G; 2645 T; 0 other;
XX
Query Match 92.0%; Score 18.4; DB 24; Length 5453;
Best Local Similarity 95.0%; Pred. NO. 1.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTTTGTGTT 20
Db 983 TTAGTTTGTGTTTGTGTT 1002
XX
RESULT 12
ABL32448
ID ABL32448 standard; DNA: 7238 BP.
XX
AC ABL32448;
XX
DT 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 421.
XX
KW Human: immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antitumour; antidiabetic; antipsoriatic;
KW antineoplastic; antidiabetic; antidiabetic; antipsoriatic;
KW antineoplastic; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200200928-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07537.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
XX

PA (EPiG-) EPIGENOMICS AG.
 XX
 PI Olek A. Plepenbrock C, Berlin K;
 XX
 DR WPI: 2002-130909/17.
 XX
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation
 XX
 PS Claim 1: SEQ ID NO 421; 32pp + Sequence Listing; German.
 XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 SQ Sequence 7238 BP; 2414 A; 84 C; 1401 G; 3339 T; 0 other;
 XX
 Query Match 92.0%; Score 18.4; DB 24; Length 7238;
 Best Local Similarity 95.0%; Pred. No. 1.2e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 TTAGTCTTAGTTTGGTTT 20
 Db 6279 TTAGTTTGTAGTTTGGTTT 6298
 XX
 RESULT 13
 ABL32072
 ID ABL32072 standard; DNA; 7353 BP.
 XX
 AC ABL32072;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Human immune system associated gene SEQ ID NO: 45.
 XX
 KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antiarteriosclerotic; antihaemic; cytosatic; nootropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antirheumatic; antiarthritic; antidiabetic; antiparasitic;
 KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200200928-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EP07537.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPiG-) EPIGENOMICS AG.
 XX
 PI Olek A. Plepenbrock C, Berlin K;
 XX
 DR WPI: 2002-130909/17.
 XX
 XX Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation
 XX
 PS Claim 1; SEQ ID NO 45; 32pp + Sequence Listing; German.

XX
 CC The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.
 XX
 SQ Sequence 7353 BP; 1544 A; 384 C; 1956 G; 3469 T; 0 other;
 XX
 Query Match 92.0%; Score 18.4; DB 24; Length 7353;
 Best Local Similarity 95.0%; Pred. No. 1.2e+02;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 TTAGTCTTAGTTTGGTTT 20
 Db 2404 TTAGTTTGTAGTTTGGTTT 2423
 XX
 RESULT 14
 AAD28362
 ID AAD28362 standard; DNA; 7353 BP.
 XX
 AC AAD28362;
 XX
 DT 22-APR-2002 (first entry)
 XX
 DE Human chemically treated genomic DNA #3.
 XX
 KW Human; cytosatic; antidepressant; neuroleptic; nootropic; antiaddictive;
 KW adrenergic alpha-1C-receptor; cytosine methylation; therapy; alcoholism;
 KW behavioural disorder; neurological; psychiatric; cancer; schizophrenia;
 KW Tourette's syndrome; smoking; human immunodeficiency virus dementia;
 KW drug abuse; migraine; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200202809-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EP07540.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPiG-) EPIGENOMICS AG.
 XX
 PI Olek A. Plepenbrock C, Berlin K;
 XX
 DR WPI: 2002-154759/20.
 XX
 XX Novel nucleic acid useful for diagnosis and therapy of behavioral
 PT disorder, neurological disorder and cancer, comprises a sequence of a
 PT segment of chemically pretreated DNA of adrenergic alpha-1C-receptor
 PT gene
 XX
 PS Claim 1; Page 32-36; 190pp; English.
 XX
 CC The invention relates to nucleic acids comprising a segment of chemically
 CC pretreated DNA of adrenergic alpha-1C-receptor gene. The invention also
 CC relates to oligonucleotides or peptide nucleic acid (PNA) oligomers
 CC useful for detecting cytosine methylations. The pretreated DNA is useful
 CC for the diagnosis or therapy of behavioural disorders, neurological
 CC disorders and cancer, in particular major depressive disorder, Tourette's
 CC syndrome, schizophrenia, psychiatric and neurological disorders, smoking,
 CC drug abuse, alcoholism, personality traits, compulsive gambling, human
 CC immunodeficiency virus dementia, migraine, behaviours in schizophrenic
 CC and schizoaffective patients, and suicidal behaviour in patients with
 CC schizophrenia. The nucleic acid is useful for detecting the methylation
 CC state of all Cpg dinucleotides and/or single nucleotide polymorphisms

CC (SNPs). The present sequence is human chemically treated genomic DNA.
XX
SQ Sequence 7353 BP; 1544 A; 384 C; 1956 G; 3469 T; 0 other;
Query Match 92.0%; Score 18.4; DB 24; Length 7353;
Best Local Similarity 95.0%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTACTTCTAGTTTGTAGTT 20
||||| |||||||||
Db 2404 TTACTTATAGTTTGTAGTT 2423
RESULT 15
ABN80085
ID ABN80085 standard; DNA; 10682 BP.
XX
AC ABN80085;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human chemically modified disease associated gene SEQ ID NO 102.
XX
KW Human; development; homeobox gene; HOX; diabetes; cancer; apoptosis;
KW heart disease; epilepsy; histone deacetylation; muscular dystrophy;
KW dwarfism; single nucleotide polymorphism; SNP; cytosine methylation;
KW antidiabetic; cytostatic; anticonvulsant; ds.
XX
OS Homo sapiens.
OS Synthetic.
PN WO200200927-A2.
XX
PD 03-JAN-2002.
XX
PF 02-JUL-2001; 2001WO-EP07536.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPiG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-130908/17.
XX
PT Novel nucleic acid useful for diagnosis and therapy of diseases
PT associated with development genes such as diabetes, comprises a
PT sequence of a segment of chemically pretreated DNA of genes associated
PT with development -
XX
PS Claim 1; SEQ ID NO 102; 27pp; English.
XX
CC The invention relates to a nucleic acid (I) comprising a sequence at
CC least 18 bases in length of a segment of chemically pretreated DNA (II)
CC of genes associated with development selected from 87 genes listed in
CC the specification such as ACCPN, ADFN, or AFDI and comprising one of 350
CC sequences (ABN79984-ABN80333) or their complements. The invention is
CC useful for the diagnosis or therapy of diseases associated with
CC development genes, in particular disease related to homeobox containing
CC genes (HOX), like diabetes, cancer, apoptosis related diseases, syndromes
CC associated with congenital heart disease, epilepsy, diseases related to
CC histone deacetylation, Currarino syndrome, diseases related with the
CC development of the brain and limb girdle muscular dystrophy and dwarfism.
CC Oligomers specific to each of the genes are useful for detecting the
CC methylation state of all CpG dinucleotides within the 350 sequences or
CC (II) and their complementary sequences, as primer oligonucleotides for
CC the amplification of the 350 sequences, (II) and/or their complements and
CC as oligomer probes for detecting the cytosine methylation state and/or
CC single nucleotide polymorphisms (SNPs).
CC Note: The sequence data for this patent did not form part of the printed
CC specification but is based on sequence information supplied to Derwent by
CC the European Patent Office.

XX
SQ Sequence 10682 BP; 3158 A; 263 C; 2318 G; 4943 T; 0 other;
Query Match 92.0%; Score 18.4; DB 24; Length 10682;
Best Local Similarity 95.0%; Pred. No. 1.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTACTTCTAGTTTGTAGTT 20
||||| |||||||||
Db 5298 TTACTTTTGTAGTTTGTAGTT 5317
Search completed: April 6, 2003, 12:05:26
Job time : 155 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:12:37 ; Search time 45.25 seconds
(without alignments)
135.548 Million cell updates/sec

Title: US-09-980-265-10

Perfect score: 20
Sequence: 1 ttagtcttagtttagtt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2-6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2-6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2-6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2-6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2-6/ptodata/2/ina/PCTUS_COMB.seq:*
6: /cgn2-6/ptodata/2/ina/Backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	16.8	84.0	1602 2	US-08-770-544-3 Sequence 3, Appli
2	15.8	79.0	362 4	US-09-605-785-442 Sequence 442, App
3	15.8	79.0	362 4	US-09-439-313-442 Sequence 442, App
4	15.8	79.0	362 4	US-09-352-616A-442 Sequence 46, Appl
5	15.8	79.0	427 4	US-09-556-877-46 Sequence 46, Appl
6	15.8	79.0	427 4	US-09-620-412C-46 Sequence 46, Appl
7	15.8	79.0	4057 4	US-08-894-997-49 Sequence 49, Appl
8	15.8	79.0	19011 4	US-08-310-356-36 Sequence 36, Appl
9	15.8	79.0	19557 5	PCT-US92-06300-1 Sequence 1, Appli
10	15.2	76.0	201 4	US-09-134-001C-2584 Sequence 2584, Ap
11	15.2	76.0	313 4	US-09-221-017B-84 Sequence 84, Appl
12	15.2	76.0	685 4	US-09-122-400B-10 Sequence 10, Appl
13	15.2	76.0	703 4	US-08-737-371A-1 Sequence 1, Appli
14	15.2	76.0	2181 5	PCT-US95-05853-1 Sequence 1, Appli
15	15.2	76.0	2181 5	PCT-US95-05853-1 Sequence 1, Appli
16	15.2	76.0	2251 4	US-08-991-677-11 Sequence 11, Appl
17	15.2	76.0	3026 4	US-09-149-476-194 Sequence 194, App
18	15.2	76.0	3054 4	US-09-149-476-194 Sequence 194, App
19	15.2	76.0	6496 4	US-09-221-017B-543 Sequence 543, App
20	15.2	76.0	18073 4	US-09-078-294-12 Sequence 12, Appl
21	15.2	76.0	59065 4	US-09-813-817-3 Sequence 3, Appli
22	15.2	76.0	59065 4	US-09-978-197-3 Sequence 3, Appli
23	15.2	76.0	246240 2	US-08-724-394A-20 Sequence 20, Appl
24	15.2	76.0	246240 2	US-08-724-394A-21 Sequence 21, Appl
25	15.2	76.0	246240 2	US-08-724-394A-22 Sequence 22, Appl
26	14.8	74.0	134 1	US-08-710-082-12 Sequence 12, Appl
27	14.8	74.0	134 3	US-08-913-462-12 Sequence 12, Appl

c 28	14.8	74.0	134 4	US-09-624-482-12 Sequence 12, Appl
29	14.8	74.0	154 3	US-08-710-082-13 Sequence 13, Appl
30	14.8	74.0	154 3	US-08-913-462-13 Sequence 13, Appl
31	14.8	74.0	154 4	US-09-624-482-13 Sequence 13, Appl
32	14.8	74.0	324 4	US-08-651-155B-22 Sequence 22, Appl
33	14.8	74.0	1180 4	US-09-276-531-51 Sequence 51, Appl
34	14.8	74.0	1260 2	US-08-747-381-1 Sequence 1, Appli
35	14.8	74.0	1260 3	US-08-969-761-1 Sequence 1, Appli
36	14.8	74.0	1260 3	US-09-225-443-1 Sequence 1, Appli
37	14.8	74.0	4973 4	US-09-381-862-6 Sequence 6, Appli
38	14.8	74.0	7278 4	US-09-091-219-1 Sequence 1, Appli
c 39	14.8	74.0	8657 4	US-08-961-527-44 Sequence 44, Appl
40	14.8	74.0	11056 4	US-09-004-838-23 Sequence 23, Appl
41	14.8	74.0	15062 4	US-09-004-838-89 Sequence 89, Appl
42	14.8	74.0	111282 4	US-09-754-250-3 Sequence 3, Appli
c 43	14.8	74.0	111282 4	US-09-754-250-3 Sequence 3, Appli
c 44	14.4	72.0	38 3	US-08-535-057A-9 Sequence 9, Appli
c 45	14.4	72.0	564 4	US-09-134-001C-2343 Sequence 2343, Ap

ALIGNMENTS

RESULT 1
US-08-770-544-3
; Sequence 3, Application US/08770544
; Patent No. 5907085
; GENERAL INFORMATION:
; APPLICANT: Gonsalves, Dennis
; APPLICANT: Ling, Kai-Siu
; TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND
; TITLE OF INVENTION: THEIR USES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,544
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60009008
; FILING DATE: 21-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/621
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1602 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-770-544-3
Query Match 84.0%; Score 16.8; DB 2; Length 1602;
Best Local Similarity 90.0%; Pred. No. 65;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Oy 1 TTAGTCTTAGTTTAGTT 20

Db 1190 TTGCTCTAGTTTACTT 1209

RESULT 2

US-09-605-785-442/C
; Sequence 442, Application US/09605785
; Patent No. 6321716
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C16
; CURRENT APPLICATION NUMBER: US/09/605,785
; CURRENT FILING DATE: 2000-06-27
; NUMBER OF SEQ ID NOS: 835
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 442
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-605-785-442

Query Match 79.0%; Score 15.8; DB 4; Length 362;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TTAGTCTCTAGTTTACTT 19

Db 346 TTAATCTTAGTTTATT 328

RESULT 3

US-09-439-313-442/C
; Sequence 442, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yugu
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Marc
; APPLICANT: Stolk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 442
; LENGTH: 362

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-439-313-442

Query Match 79.0%; Score 15.8; DB 4; Length 362;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TTAGTCTCTAGTTTACTT 19

Db 346 TTAATCTTAGTTTATT 328

RESULT 4

US-09-352-616A-442/C
; Sequence 442, Application US/09352616A
; Patent No. 6395278
; GENERAL INFORMATION:
; APPLICANT: Dillon, Davin C.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang, Yugu
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer Lynn
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.427C8
; CURRENT APPLICATION NUMBER: US/09/352,616A
; CURRENT FILING DATE: 1999-07-13
; NUMBER OF SEQ ID NOS: 472
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 442
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-352-616A-442

Query Match 79.0%; Score 15.8; DB 4; Length 362;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TTAGTCTCTAGTTTACTT 19

Db 346 TTAATCTTAGTTTATT 328

RESULT 5

US-09-556-877-46
; Sequence 46, Application US/09556877
; Patent No. 6432916
; GENERAL INFORMATION:
; APPLICANT: Probst, Peter
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir
; APPLICANT: Fling, Steve
; APPLICANT: Maisonneuve, Jeff
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C5
; CURRENT APPLICATION NUMBER: US/09/556,877
; CURRENT FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 305
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 46
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n-A,T,C or G
US-09-556-877-46

Query Match 79.0%; Score 15.8; DB 4; Length 427;

Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTACT 19
|| |||||
DB 30 TTTTCTTAGTTTACT 48

RESULT 6
US-09-620-412C-46
; Sequence 46, Application US/09620412C
; Patent No. 6448234
; GENERAL INFORMATION:
; APPLICANT: Steven P. Fling
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C7
; CURRENT APPLICATION NUMBER: US/09/620,412C
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 46
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n-A,T,C or G
US-09-620-412C-46

Query Match 79.0%; Score 15.8; DB 4; Length 427;
Best Local Similarity 89.5%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTACT 19
|| |||||
DB 30 TTTTCTTAGTTTACT 48

RESULT 7
US-08-894-997-49
; Sequence 49, Application US/08894997A
; Patent No. 6270990
; GENERAL INFORMATION:
; APPLICANT: Schenbert, David J
; TITLE OF INVENTION: NEURON-RESTRICTIVE SILENCER FACTOR
; FILE REFERENCE: 17810-502 NRSE
; CURRENT APPLICATION NUMBER: US/08/894,997A
; EARLIER FILING DATE: 1998-01-06
; EARLIER APPLICATION NUMBER: PCT/US96/02817
; EARLIER FILING DATE: 1996-03-01
; EARLIER APPLICATION NUMBER: 08/398,590
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 49
; LENGTH: 4057
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: gene
; LOCATION: (1)-(4057)
; OTHER INFORMATION: Human NSRF
US-08-894-997-49

Query Match 79.0%; Score 15.8; DB 4; Length 4057;
Best Local Similarity 89.5%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 TTAGTCTTAGTTTACT 20
|| ||||| || |||||

DB 3581 TAGTCTTAATTTTACT 3599

RESULT 8
US-08-310-356-36/C
; Sequence 36, Application US/08310356
; Patent No. 5648243
; GENERAL INFORMATION:
; APPLICANT: Hurwitz, David R
; APPLICANT: Nathan, Margaret
; APPLICANT: Shaul, Moshe
; TITLE OF INVENTION: Transgenic Protein Production
; NUMBER OF SEQUENCES: 36
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Rhone-Poulenc Rorer Legal Department
; STREET: 500 Arcola Road
; CITY: Collegeville
; STATE: PA
; COUNTRY: USA
; ZIP: 19426
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: Macintosh
; OPERATING SYSTEM: Macintosh System 7.0
; SOFTWARE: Microsoft Word Version 5.0 (Patentin)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/310,356
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/737,853
; FILING DATE: 31-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Goodman, Rosanne
; REGISTRATION NUMBER: 32,534
; REFERENCE/DOCKET NUMBER: A0856
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 454-3817
; TELEFAX: (215) 454-3808
; INFORMATION FOR SEQ ID NO: 36:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1737..1854
; OTHER INFORMATION: /number= 1
; OTHER INFORMATION: /citation= (11)
; FEATURE:
; NAME/KEY: Intron
; LOCATION: 1855..2563
; OTHER INFORMATION: /number= 1
; OTHER INFORMATION: /label= Intron1
; OTHER INFORMATION: /citation= (11)
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2564..2621
; OTHER INFORMATION: /number= 2
; OTHER INFORMATION: /citation= (11)
; FEATURE:
; NAME/KEY: Intron
; LOCATION: 2622..4075
; OTHER INFORMATION: /number= 2
; OTHER INFORMATION: /citation= (11)
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4076..4208
; OTHER INFORMATION: /number= 3
; OTHER INFORMATION: /citation= (11)
; FEATURE:
; NAME/KEY: Intron

LOCATION: 4209..6040
OTHER INFORMATION: /number= 3
FEATURE:
NAME/KEY: exon
LOCATION: 6041..6252
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: intron
LOCATION: 6253..6801
OTHER INFORMATION: /number= 4
FEATURE:
NAME/KEY: exon
LOCATION: 6802..6934
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: intron
LOCATION: 6935..7758
OTHER INFORMATION: /number= 5
FEATURE:
NAME/KEY: exon
LOCATION: 7759..7856
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: intron
LOCATION: 7857..9443
OTHER INFORMATION: /number= 6
FEATURE:
NAME/KEY: exon
LOCATION: 9444..9573
OTHER INFORMATION: /number= 7
FEATURE:
NAME/KEY: intron
LOCATION: 9574..10866
OTHER INFORMATION: /number= 7
FEATURE:
NAME/KEY: exon
LOCATION: 10867..11081
OTHER INFORMATION: /number= 8
FEATURE:
NAME/KEY: intron
LOCATION: 11082..12480
OTHER INFORMATION: /number= 8
FEATURE:
NAME/KEY: exon
LOCATION: 12481..12613
OTHER INFORMATION: /number= 9
FEATURE:
NAME/KEY: intron
LOCATION: 12614..13701
OTHER INFORMATION: /number= 9
FEATURE:
NAME/KEY: exon
LOCATION: 13702..13799
OTHER INFORMATION: /number= 10
FEATURE:
NAME/KEY: intron
LOCATION: 13800..14976
OTHER INFORMATION: /number= 10
FEATURE:
NAME/KEY: exon
LOCATION: 14977..15115
OTHER INFORMATION: /number= 11
FEATURE:
NAME/KEY: intron
LOCATION: 15116..15533
OTHER INFORMATION: /number= 11
FEATURE:
NAME/KEY: exon
LOCATION: 15534..15757
OTHER INFORMATION: /number= 12
FEATURE:
NAME/KEY: intron
LOCATION: 15758..16949

OTHER INFORMATION: /number= 12
FEATURE:
NAME/KEY: exon
LOCATION: 16950..17082
OTHER INFORMATION: /number= 13
FEATURE:
NAME/KEY: intron
LOCATION: 17083..17696
OTHER INFORMATION: /number= 13
FEATURE:
NAME/KEY: exon
LOCATION: 17697..17764
OTHER INFORMATION: /number= 14
FEATURE:
NAME/KEY: intron
LOCATION: 17765..18534
OTHER INFORMATION: /number= 14
FEATURE:
NAME/KEY: CDS
LOCATION: join(1776..1854, 2564..2621, 4076..4208,
LOCATION: 6041..6252, 6802..6934, 7759..7856, 9444..9573,
LOCATION: 10867..11081, 12481..12613, 13702..13799,
LOCATION: 14977..15115, 15534..15757, 16950..17082,
LOCATION: 17697..17741)
OTHER INFORMATION: /product= "human serum albumin"
OTHER INFORMATION: /citation= (11)
FEATURE:
NAME/KEY: exon
LOCATION: 18535..18697
OTHER INFORMATION: /number= 15
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 17742..18697
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1737..1775
PUBLICATION INFORMATION:
AUTHORS: Minghetti, P P
AUTHORS: Ruffner, D E
AUTHORS: Kuang, W-J
AUTHORS: Dennison, O E
AUTHORS: Hawkins, J W
AUTHORS: Beattie, W G
AUTHORS: Dugaiczky, A
TITLE: MOLECULAR STRUCTURE OF THE HUMAN ALBUMIN
TITLE: GENE IS REVEALED BY NUCLEOTIDE SEQUENCE WITHIN
JOURNAL: J. Biol. Chem.
VOLUME: 261
PAGES: 6747-6757
DATE: 1986
RELEVANT RESIDUES IN SEQ ID NO: 36: FROM 1 TO 19011
US-08-310-356-36
Query Match 79.0%; Score 15.8; DB 1; Length 19011;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY 1 TTAGTCTTAGTCTTTTACT 19
||||||| ||| |||||
Db 1409 TTAGTCTCAGTGTTAGT 1391
RESULT 9
PCT-US92-06300-1/c
; Sequence 1, Application PC/TUS9206300
; GENERAL INFORMATION:
; APPLICANT: Hurwitz, David R
; APPLICANT: Nathan, Margaret
; APPLICANT: Shani, Moshe
; TITLE OF INVENTION: Transgenic Protein Production
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Rhone-Poulenc Rorer, Inc.
STREET: 500 Virginia Ave., Bldg. 3A
CITY: Ft. Washington
STATE: Pennsylvania
COUNTRY: USA
ZIP: 19034
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/06300
FILING DATE: 19920730
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Goodman, Rosanne
REGISTRATION NUMBER: 52,534
REFERENCE/DOCKET NUMBER: A0856-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 962-4130
TELEFAX: (215) 962-4107
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 19557 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PUBLICATION INFORMATION:
AUTHORS: Minibetli, P P
AUTHORS: Ruffner, D E
AUTHORS: Kuang, W.-J.
AUTHORS: Demnlson, O E
AUTHORS: Hawkins, J W
AUTHORS: Beattie, W G
AUTHORS: Dugalczyk, A
TITLE: MOLECULAR STRUCTURE OF THE HUMAN ALBUMIN GENE
TITLE: IS REVEALED BY NUCLEOTIDE SEQUENCE WITHIN Q11-22
TITLE: OF CHROMOSOME 4
JOURNAL: J. Biol. Chem.
VOLUME: 261
PAGES: 6747-6757
DATE: 1986
RELEVANT RESIDUES IN SEQ ID NO: 1: FROM 1 TO 19002
PCT-US92-06300-1
Query Match 79.0%; Score 15.8; DB 5; Length 19557;
Best Local Similarity 89.5%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 TTAGTCTTACTTTTACT 19
||||| ||| |||||
Db 1964 TTAGTCTCAGGTGTTAGT 1946
RESULT 10
US-09-134-001C-2584/c
Sequence 2584, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14

NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 2584
LENGTH: 201
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-2584
Query Match 76.0%; Score 15.2; DB 4; Length 201;
Best Local Similarity 85.0%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 TTAGTCTTACTTTTACT 20
||||| ||| |||||
Db 115 TTCGTTCTTCTTTTACT 96
RESULT 11
US-09-221-017B-84/c
Sequence 84, Application US/09221017B
Patent No. 644799
GENERAL INFORMATION:
APPLICANT: Ross, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & ROERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FASTSEQ for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Monroy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 84:
SEQUENCE CHARACTERISTICS:
LENGTH: 313 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature

LOCATION: 1...313
US-09-221-017B-84
Query Match 76.0%; Score 15.2; DB 4; Length 313;
Best Local Similarity 85.0%; Pred. No. 3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TTAGTCTAGTTTGT 20
DB 258 TTATTTCTGTGTTTGT 239
RESULT 12
US-09-122-400B-10/c
Sequence 10, Application US/09122400B
Patent No. 6245974
GENERAL INFORMATION:
APPLICANT: Michalowski, Susan
TITLE OF INVENTION: MATRIX ATTACHMENT REGIONS
FILE REFERENCE: Michalowski and Spiker
CURRENT APPLICATION NUMBER: US/09/122,400B
CURRENT FILING DATE: 1998-07-24
PRIOR APPLICATION NUMBER: 60/066,118
PRIOR FILING DATE: 1997-08-06
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 10
LENGTH: 685
TYPE: DNA
ORGANISM: Nicotiana tabacum
US-09-122-400B-10
Query Match 76.0%; Score 15.2; DB 4; Length 685;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TTAGTCTAGTTTGT 20
DB 636 TTAGTTTATTTTATTT 617
RESULT 13
US-09-221-017B-598/c
Sequence 598, Application US/09221017B
Patent No. 6444799
GENERAL INFORMATION:
APPLICANT: Ross, Bruce C.
TITLE OF INVENTION: P. GINGIVALIS NUCLEOTIDES AND USES THEREOF
NUMBER OF SEQUENCES: 1120
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 755 PAGE MILL ROAD
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: Windows
SOFTWARE: FastSeq for Windows Version 2.0b
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/221,017B
FILING DATE: 23-DEC-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1182
FILING DATE: 31-DEC-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: P1546
FILING DATE: 30-JAN-1998
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PP2911
FILING DATE: 09-APR-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU98/01023
FILING DATE: 10-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Monroy, Gladys H
REGISTRATION NUMBER: 32,430
REFERENCE/DOCKET NUMBER: 27340-20021.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-813-5600
TELEFAX: 650-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 598:
SEQUENCE CHARACTERISTICS:
LENGTH: 703 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: UNKNOWN
ORIGINAL SOURCE:
ORGANISM: PORPHYROMONAS GINGIVALIS
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1...703
US-09-221-017B-598
Query Match 76.0%; Score 15.2; DB 4; Length 703;
Best Local Similarity 85.0%; Pred. No. 2.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TTAGTCTAGTTTGT 20
DB 200 TTATTTCTGTGTTTGT 181
RESULT 14
US-08-737-371A-1/c
Sequence 1, Application US/08737371A
Patent No. 5959094
GENERAL INFORMATION:
APPLICANT: David WALLACH
APPLICANT: Peter KUHNERT
APPLICANT: Gotz EHRHARDT
APPLICANT: Oliver KEMPER
TITLE OF INVENTION: p75 TNF RECEPTOR PROMOTERS
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,371A
FILING DATE: 08-NOVEMBER-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05853
FILING DATE: 11-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 109,633
FILING DATE: 11-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618

REFERENCE/DOCKET NUMBER: WALLACH-14
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2181 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2094..2171
US-08-737-371A-1

Query Match 76.0%; Score 15.2; DB 2; length 2181;
Best Local Similarity 85.0%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 TTAGTCTTAGTCTTAGT 20
||| ||||| ||||| ||
Db 1591 TTATCTTATTTTAA 1572

RESULT 15

PCT-US95-05853-1/c
Sequence 1, Application PC/TUS9505853
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: p75 TNF RECEPTOR PROMOTERS
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05853
FILING DATE: 11-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: IL 109,633
FILING DATE: 11-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: WALLACH-14 PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2181 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 2094..2171
PCT-US95-05853-1

Query Match 76.0%; Score 15.2; DB 5; length 2181;
Best Local Similarity 85.0%; Pred. No. 2.8e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 TTAGTCTTAGTCTTAGT 20
||| ||||| ||||| ||
Db 1591 TTATCTTATTTTAA 1572

Search completed: April 6, 2003, 11:55:45
Job time : 53.25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:24:32 ; Search time 58 Seconds
(without alignments)
302.472 Million cell updates/sec

Title: US-09-980-265-10

Perfect score: 20

Sequence: 1 ttactcttactttttagt 20

Scoring table: IDENTITY_NUC

GAPD 10.0, Gapext 1.0

Searched: 593429 seqs, 438583890 residues 1186858

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_MA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PC1_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PC1US_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	16.8	84.0	548	10 US-09-924-035A-396	Sequence 396, App
2	16.8	84.0	63000	10 US-09-780-172-18	Sequence 18, Appl
3	16.4	82.0	50000	9 US-10-060-763-4	Sequence 4, Appl
4	16.4	82.0	50000	12 US-10-063-763-4	Sequence 4, Appl
5	16	80.0	219	10 US-09-998-588-2498	Sequence 2498, App
c 6	15.8	79.0	362	9 US-10-012-896-442	Sequence 442, App
c 7	15.8	79.0	362	9 US-09-895-793-442	Sequence 442, App
c 8	15.8	79.0	362	9 US-09-895-793-442	Sequence 442, App
c 9	15.8	79.0	362	10 US-09-880-107-2657	Sequence 2657, App
c 10	15.8	79.0	362	10 US-09-880-107-2657	Sequence 2657, App
c 11	15.8	79.0	362	10 US-09-880-107-2657	Sequence 2657, App
c 12	15.8	79.0	427	10 US-09-841-132-46	Sequence 46, Appl
c 13	15.8	79.0	427	10 US-09-841-132-46	Sequence 46, Appl
c 14	15.8	79.0	579	10 US-09-841-132-46	Sequence 579, App
c 15	15.8	79.0	24132	10 US-09-764-860-661	Sequence 661, App
c 16	15.8	79.0	513509	9 US-09-754-853A-4	Sequence 4, Appl
c 17	15.8	79.0	640681	10 US-09-790-988-1	Sequence 1, Appl
c 18	15.4	77.0	269	10 US-09-822-849A-138	Sequence 138, App
c 19	15.4	77.0	1505	10 US-09-822-849A-138	Sequence 138, App

20	15.4	77.0	2000	9 US-09-938-842A-4209	Sequence 4209, App
21	15.4	77.0	659158	9 US-09-771-208-20	Sequence 20, Appl
22	15.2	76.0	192	10 US-09-815-242-2912	Sequence 2912, App
23	15.2	76.0	192	10 US-09-815-242-2918	Sequence 2918, App
24	15.2	76.0	192	10 US-09-815-242-2947	Sequence 2947, App
25	15.2	76.0	192	10 US-09-815-242-3417	Sequence 3417, App
26	15.2	76.0	192	10 US-09-815-242-3488	Sequence 3488, App
c 27	15.2	76.0	256	10 US-09-783-590-9439	Sequence 9439, App
c 28	15.2	76.0	368	9 US-09-232-785-352	Sequence 352, App
c 29	15.2	76.0	378	10 US-09-815-343-1089	Sequence 1089, App
c 30	15.2	76.0	397	10 US-09-867-701-5873	Sequence 5873, App
c 31	15.2	76.0	424	9 US-09-796-692-4100	Sequence 4100, App
c 32	15.2	76.0	462	10 US-09-960-352-7660	Sequence 7660, App
c 33	15.2	76.0	494	10 US-09-867-701-6731	Sequence 6731, App
c 34	15.2	76.0	572	10 US-09-864-761-12035	Sequence 12035, App
c 35	15.2	76.0	620	9 US-09-908-931B-37	Sequence 37, Appl
c 36	15.2	76.0	667	9 US-09-908-931B-12	Sequence 12, Appl
c 37	15.2	76.0	685	10 US-09-816-894-10	Sequence 10, Appl
c 38	15.2	76.0	697	10 US-09-770-149-277	Sequence 277, App
c 39	15.2	76.0	701	10 US-09-770-149-251	Sequence 251, App
c 40	15.2	76.0	760	9 US-10-040-739-1332	Sequence 1332, App
c 41	15.2	76.0	1019	9 US-10-278-173-153	Sequence 153, App
c 42	15.2	76.0	1070	9 US-10-278-173-133	Sequence 133, App
c 43	15.2	76.0	1202	10 US-09-822-830A-591	Sequence 591, App
c 44	15.2	76.0	1557	9 US-10-001-873-1	Sequence 1, Appl
c 45	15.2	76.0	1707	9 US-10-108-605-152	Sequence 152, App

ALIGNMENTS

RESULT 1
US-09-924-035A-396/c
Sequence 396, Application US/09924035A
Patent No. US20020142319A1
GENERAL INFORMATION:
APPLICANT: Glaxo, Jm
TITLE OF INVENTION: Expressed Sequences of Arabidopsis
FILE REFERENCE: 2011US
CURRENT APPLICATION NUMBER: US/09/924, 035A
PRIOR APPLICATION NUMBER: 2000-08-11
PRIOR FILING DATE: 1999-08-13
NUMBER OF SEQ ID NOS: 900
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 396
LENGTH: 548
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)...(548)
OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-396

Query Match 84.0%; Score 16.8; DB 10;
Best Local Similarity 90.0%; Pred. No. 3,4e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 172 TTACTCTTACTTGTAGTT 153
172 TTACTCTTACTTGTAGTT 153

RESULT 2
US-09-780-172-18
Sequence 18, Application US/09780172
Patent No. US20020147163A1
GENERAL INFORMATION:
APPLICANT: Robert McKay
APPLICANT: Susan M. Freiler
APPLICANT: Jacqueline Wyatt

```

; TITLE OF INVENTION: ANTISENSE MODULATION OF CASEIN KINASE 2-ALPHA EXPRESSION
; FILE REFERENCE: RTS-0159
; CURRENT APPLICATION NUMBER: US/09/780,172
; CURRENT FILING DATE: 2001-02-08
; NUMBER OF SEQ ID NOS: 96
; SEQ ID NO 18
; LENGTH: 63000
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
US-09-780-172-18

Query Match      84.0%; Score 16.8; DB 10; Length 63000;
Best Local Similarity 90.0%; Pred. No. 5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTACTCTTAGTTTGTAGT 20
Db 18859 TTAGTCTTAGTTTGTAGT 18878

RESULT 3
US-10-060-763-4
; Sequence 4, Application US/10060763
; Publication No. US20030022286A1
; GENERAL INFORMATION:
; APPLICANT: Curtlis, Rory A.J.
; TITLE OF INVENTION: No. US20030022286A1el Transporter-Like Genes and Uses Therefor
; FILE REFERENCE: 10147-8
; CURRENT APPLICATION NUMBER: US/10/060,763
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-060-763-4

Query Match      82.0%; Score 16.4; DB 9; Length 50000;
Best Local Similarity 94.4%; Pred. No. 7.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTACTCTTAGTTTGTAG 18
Db 4192 TTACTCTTAGTTTGTAG 4209

RESULT 4
US-10-063-763-4
; Sequence 4, Application US/10063763
; Patent No. US20010001663A1
; GENERAL INFORMATION:
; APPLICANT: Curtlis, Rory A.J.
; TITLE OF INVENTION: No. US20010001663A1el Transporter-Like Genes and Uses Therefor
; FILE REFERENCE: 10147-8
; CURRENT APPLICATION NUMBER: US/10/063,763
; CURRENT FILING DATE: 2002-01-30
; PRIOR APPLICATION NUMBER: 09/365,162
; PRIOR FILING DATE: 1999-JUL-30
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 50000
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-063-763-4

Query Match      82.0%; Score 16.4; DB 12; Length 50000;
Best Local Similarity 94.4%; Pred. No. 7.1e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTTTGTAG 18
```

```

Db 4192 TTACTCTTAGTTTGTAG 4209
|||||
RESULT 5
US-09-998-598-2498
; Sequence 2498, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Chenault, Ruth A.
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 2498
; LENGTH: 219
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-598-2498

Query Match      80.0%; Score 16; DB 10; Length 219;
Best Local Similarity 100.0%; Pred. No. 6.6e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 TTCTTAGTTTGTAGT 20
Db 143 TTCTTAGTTTGTAGT 158

RESULT 6
US-10-012-896-442/c
; Sequence 442, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqin
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hurral, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Roy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Mantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 442
; LENGTH: 362
```

```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-442

Query Match          79.0%; Score 15.8; DB 9; Length 362;
Best Local Similarity 89.5%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 TTAGTCTTAGTCTTTAGT 19
      ||| ||||| ||||| |||
Db      346 TTAATCTTAGTCTTTTATT 328

RESULT 7
US-09-895-793-442/c
; Sequence 442, Application US/09895793
; Publication No. US20020192763A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hurai, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C2
; CURRENT APPLICATION NUMBER: US/09/895,793
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 442
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-793-442

Query Match          79.0%; Score 15.8; DB 9; Length 362;
Best Local Similarity 89.5%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 TTAGTCTTAGTCTTTAGT 19
      ||| ||||| ||||| |||
Db      346 TTAATCTTAGTCTTTTATT 328

RESULT 8
US-09-895-814-442/c
; Sequence 442, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Kalos, Michael D.
```

```
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hurai, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 442
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-442

Query Match          79.0%; Score 15.8; DB 9; Length 362;
Best Local Similarity 89.5%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY      1 TTACTCTTAGTCTTTTACT 19
      ||| ||||| ||||| |||
Db      346 TTAATCTTAGTCTTTTATT 328

RESULT 9
US-09-759-143-442/c
; Sequence 442, Application US/09759143
; Patent No. US20020022248A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yugu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel
; APPLICANT: Wang, Aljun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C23
; CURRENT APPLICATION NUMBER: US/09/759,143
; CURRENT FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 442
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-759-143-442
```

Query Match 79.0%; Score 15.8; DB 10; Length 362;
Best Local Similarity 89.5%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTTTACT 19
||| ||||| ||||| |||
Db 346 TTAATCTTACTTTTATT 328

RESULT 10
US-09-780-669-442/C
; Sequence 442, Application US/09780669
; Patent No. US20020051977A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqi
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Wang, Aijun
; APPLICANT: Li, Samuel
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghon, Raymond L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.427C24
; CURRENT APPLICATION NUMBER: US/09/780.669
; CURRENT FILING DATE: 2001-02-09
; NUMBER OF SEQ ID NOS: 943
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 442
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-780-669-442

Query Match 79.0%; Score 15.8; DB 10; Length 362;
Best Local Similarity 89.5%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTTTACT 19
||| ||||| ||||| |||
Db 346 TTAATCTTACTTTTATT 328

RESULT 11
US-09-822-827-442/C
; Sequence 442, Application US/09822827
; Patent No. US20020081680A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.534C1
; CURRENT APPLICATION NUMBER: US/09/822.827
; CURRENT FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 982
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 442
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-822-827-442

Query Match 79.0%; Score 15.8; DB 10; Length 362;
Best Local Similarity 89.5%; Pred. No. 8.3e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTTTACT 19
||| ||||| ||||| |||
Db 346 TTAATCTTACTTTTATT 328

RESULT 12
US-09-841-132-46
; Sequence 46, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
; APPLICANT: Bhatia, Ajay
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Probst, Peter
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
; FILE REFERENCE: 210121.469C8
; CURRENT APPLICATION NUMBER: US/09/841.132
; CURRENT FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 599
; SOFTWARE: FastSeq for Windows Version 3.0/4.0
; SEQ ID NO 46
; LENGTH: 427
; TYPE: DNA
; ORGANISM: Chlamydia
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (20)
; OTHER INFORMATION: n=A,T,C or G
US-09-841-132-46

Query Match 79.0%; Score 15.8; DB 10; Length 427;
Best Local Similarity 89.5%; Pred. No. 8.4e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTTTTACT 19
||| ||||| ||||| |||
Db 30 TTTTCTTACTTTTACT 48

RESULT 13
US-09-880-107-2657/C
; Sequence 2657, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-NO
; CURRENT APPLICATION NUMBER: US/09/880.107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2657
; LENGTH: 442
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 N70481
US-09-880-107-2657

Query Match 79.0%; Score 15.8; DB 10; Length 442;

Best Local Similarity 89.5%; Pred. No. 8.4e+02;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Oy 2 TAGTCTTAGTTTGTAGTT 20
 |||||||
 Db 103 TAGTCTTAATTATT 85

RESULT 14

US-09-841-132-546/C
 ; Sequence 546, Application US/09841132
 ; Patent No. US20020061848A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bhatia, Ajay
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Probst, Peter
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
 ; FILE REFERENCE: 210121.469C8
 ; CURRENT APPLICATION NUMBER: US/09/841.132
 ; CURRENT FILING DATE: 2001-04-23
 ; NUMBER OF SEQ ID NOS: 599
 ; SOFTWARE: FastSeq for Windows Version 3.0/4.0
 ; SEQ ID NO 546
 ; LENGTH: 579
 ; TYPE: DNA
 ; ORGANISM: C. Trachomatis D serovar
 US-09-841-132-546

Query Match

Best Local Similarity 79.0%; Score 15.8; DB 10; Length 579;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 1 TTAGTCTTAGTTTGTAGT 19
 |||||||
 Db 579 TTTTCTTAGTTTGTAGT 561

RESULT 15

US-09-764-860-661
 ; Sequence 661, Application US/09764860
 ; Patent No. US20020094953A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PC008
 ; CURRENT APPLICATION NUMBER: US/09/764.860
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1198
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 661
 ; LENGTH: 24132
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-764-860-661

Query Match

Best Local Similarity 79.0%; Score 15.8; DB 10; Length 24132;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 2 TAGTCTTAGTTTGTAGTT 20
 |||||||
 Db 13693 TAGTCTTAATTATT 13711

Search completed: April 6, 2003, 12:09:57
 Job time : 86 secs

THIS PAGE BLANK (USPTO)

JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 155877)
MORLEY, K.C.
Direct Submission
Submitted (19-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GYWL
Center clone name: CH230-117D11
----- Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100K of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 73163 bases at least Q40
Consensus quality: 82138 bases at least Q30
Consensus quality: 87186 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 85 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1062: contig of 1062 bp in length
1063 1162: gap of unknown length
1163 2171: contig of 1009 bp in length
2172 2271: gap of unknown length
2271 3321: contig of 1050 bp in length
3321 3421: gap of unknown length
3421 4463: contig of 1042 bp in length
4463 4564: gap of unknown length
4564 5587: contig of 1024 bp in length
5587 5687: gap of unknown length
5687 6712: contig of 1025 bp in length
6712 6812: gap of unknown length
6812 7823: contig of 1010 bp in length
7823 7923: gap of unknown length
7923 9753: contig of 1831 bp in length
9753 9854: gap of unknown length
9854 10887: contig of 1034 bp in length
10887 10988: gap of unknown length
10988 12401: contig of 1414 bp in length
12401 12501: gap of unknown length
12501 14037: contig of 1536 bp in length
14037 14137: gap of unknown length
14137 15197: contig of 1060 bp in length
15197 15297: gap of unknown length
15297 16644: contig of 1347 bp in length
16644 16744: gap of unknown length
16744 18177: contig of 1433 bp in length
18177 18278: gap of unknown length
18278 19658: contig of 1381 bp in length
19658 19758: gap of unknown length
19758 20806: contig of 1048 bp in length
20806 20906: gap of unknown length
20906 22312: contig of 1406 bp in length
22312 22412: gap of unknown length
22412 23922: contig of 1510 bp in length
23922 24022: gap of unknown length
24022 25433: contig of 1411 bp in length
25433 25533: gap of unknown length
25533 26809: contig of 1276 bp in length
26809 26909: gap of unknown length

26910 28404: contig of 1495 bp in length
28404 28405: gap of unknown length
28405 28505: contig of 1166 bp in length
28505 29671: gap of unknown length
29671 29771: contig of 1224 bp in length
29771 30995: gap of unknown length
30995 31095: contig of 1581 bp in length
31095 32675: gap of unknown length
32675 32776: contig of 1355 bp in length
32776 34131: gap of unknown length
34131 34231: contig of 1259 bp in length
34231 35489: gap of unknown length
35489 35589: gap of unknown length
35589 37026: contig of 1437 bp in length
37026 37126: gap of unknown length
37126 38796: contig of 1670 bp in length
38796 38896: gap of unknown length
38896 40001: contig of 1105 bp in length
40001 40101: gap of unknown length
40101 41115: contig of 1014 bp in length
41115 41215: gap of unknown length
41215 42569: contig of 1354 bp in length
42569 42669: gap of unknown length
42669 43830: contig of 1161 bp in length
43830 43930: gap of unknown length
43930 45281: contig of 1351 bp in length
45281 45381: gap of unknown length
45381 47441: contig of 2060 bp in length
47441 47541: gap of unknown length
47541 49026: contig of 1485 bp in length
49026 49126: gap of unknown length
49126 50263: contig of 1137 bp in length
50263 50363: gap of unknown length
50363 52441: contig of 2078 bp in length
52441 52541: gap of unknown length
52541 53981: contig of 1440 bp in length
53981 54081: gap of unknown length
54081 55502: contig of 1421 bp in length
55502 55602: gap of unknown length
55602 56893: contig of 1291 bp in length
56893 56993: gap of unknown length
56993 58954: contig of 1960 bp in length
58954 59054: gap of unknown length
59054 60351: contig of 1298 bp in length
60351 60451: gap of unknown length
60451 61848: contig of 1397 bp in length
61848 61948: gap of unknown length
61948 64132: contig of 2184 bp in length
64132 64232: gap of unknown length
64232 65724: contig of 1392 bp in length
65724 65825: gap of unknown length
65825 66795: contig of 1071 bp in length
66795 66895: gap of unknown length
66895 68161: contig of 1266 bp in length
68161 68261: gap of unknown length
68261 69390: contig of 1129 bp in length
69390 69490: gap of unknown length
69490 70785: contig of 1295 bp in length
70785 70885: gap of unknown length
70885 73269: contig of 2384 bp in length
73269 73369: gap of unknown length
73369 75050: contig of 1681 bp in length
75050 75150: gap of unknown length
75150 76285: contig of 1135 bp in length
76285 76385: gap of unknown length
76385 78166: contig of 1781 bp in length
78166 78266: gap of unknown length
78266 80185: contig of 1919 bp in length
80185 80285: gap of unknown length
80285 81681: contig of 1396 bp in length
81681 81781: gap of unknown length
81781 83340: contig of 1559 bp in length
83340 83440: gap of unknown length
83440 85436: contig of 1996 bp in length

```

* 85437 85536: gap of unknown length

Query Match      100.0%: Score 20; DB 1; Length 155877;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTCTTAGT 20
    |||||
Db 46393 TTAGTCTTAGTCTTAGT 46412

RESULT 2
AC015120 19997 bp DNA linear HTG 16-NOV-1999
LOCUS AC015120
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
ACCESSION AC015120
VERSION AC015120.1 GI:6436215
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
REFERENCE 1 (bases 1 to 19997)
AUTHORS Adams, M. and Venter, J.C.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
COMMENT This sequence was identified as CDM:10211942 by the submitter.
For further information on this sequence e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
    source
        1..19997
        /organism="Drosophila melanogaster"
        /db_xref="taxon:7227"
BASE COUNT 5276 a 4312 c 4203 g 6206 t
ORIGIN
Query Match      95.0%: Score 19; DB 1; Length 19997;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTCTTAGT 19
    |||||
Db 1511 TTAGTCTTAGTCTTAGT 1529

RESULT 3
AC010951/c 148208 bp DNA linear HTG 01-APR-2000
LOCUS AC010951
DEFINITION Homo sapiens clone RP11-3N19, WORKING DRAFT SEQUENCE, 8 unordered
pieces.
ACCESSION AC010951
VERSION AC010951.4 GI:7341835
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 148208)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens, clone RP11-3N19
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 148208)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Bokhalekar, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collumore, A.,
Cooke, P., Dearrellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,

```

```

Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Johnson, R., Jones, C., Kann, L., Karats, A., Klein, J.,
Lewoczky, J., Lien, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., Mcgurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, F., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Strange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tessaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, D., Wu, X.,
Wyman, D., Ye, W.J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (28-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 30, 2000 this sequence version replaced gi:6087956.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information
Center project name: L2785
Center clone name: 3_N_19

----- Summary Statistics
Sequencing vector: M13; M7815: 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 142002 bases at least Q40
Consensus quality: 144771 bases at least Q30
Consensus quality: 145833 bases at least Q20
Insert size: 147000; agarose-fp
Insert size: 147508; sum-of-contigs
Quality coverage: 4.9 in Q20 bases; agarose-fp
Quality coverage: 4.9 in Q20 bases; sum-of-contigs

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2007: contig of 2007 bp in length
* 2008 2107: gap of 100 bp
* 2108 4935: contig of 2828 bp in length
* 4936 5035: gap of 100 bp
* 5036 18397: contig of 13362 bp in length
* 18398 18497: gap of 100 bp
* 18498 33384: contig of 14887 bp in length
* 33385 33484: gap of 100 bp
* 33485 46070: contig of 12586 bp in length
* 46071 46170: gap of 100 bp
* 46171 66736: contig of 20566 bp in length
* 66737 66836: gap of 100 bp
* 66837 105893: contig of 39057 bp in length
* 105894 105993: gap of 100 bp
* 105994 148208: contig of 42215 bp in length.

FEATURES
    source
        1..148208
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="RP11-3N19"
        /clone_1lb="RP11-11 Human Male BAC"
        1..2007
        /note="assembly_fragment"
        2108..4935
        /note="assembly_fragment"
        5036..18397
        /note="assembly_fragment"
        clone_end:T7
        vector_side:right"

```



```

source
1. 173471
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="RP23-251G19"
/clone_lib="RPC1-23 Female Mouse BAC"
1. 1586
/misc_feature
/note="assembly-fragment"
1687. 43654
/misc_feature
/note="assembly-fragment"
43755. 47502
/misc_feature
/note="assembly-fragment"
47603. 51368
/misc_feature
/note="assembly-fragment"
51469. 57076
/misc_feature
/note="assembly-fragment"
57177. 61964
/misc_feature
/note="assembly-fragment"
62055. 73713
/misc_feature
/note="assembly-fragment"
73816. 87823
/misc_feature
/note="assembly-fragment"
87924. 99375
/misc_feature
/note="assembly-fragment"
99476. 121365
/misc_feature
/note="assembly-fragment"
121466. 146521
/misc_feature
/note="assembly-fragment"
146622. 172750
/misc_feature
/note="assembly-fragment"
172851. 173471
/misc_feature
/note="assembly-fragment"
clone_end:77
vector_side:right"

BASE COUNT 55827 a 33466 c 33802 g 49176 t 1200 others
ORIGIN

Query Match 95.0%; Score 19; DB 1; Length 173471;
Best Local Similarity 100.0%; Pred. No. 8.8;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TAGTCTTAGTCTTAGT 20
|||||
Db 112559 TAGTCTTAGTCTTAGT 112577

RESULT 5
AC130415 186581 bp DNA linear HTG 10-AUG-2002
DEFINITION Homo sapiens chromosome 12 clone RP11-90C1, *** SEQUENCING IN
ACCESSION AC130415
VERSION AC130415.1 GI:22203172
KEYWORDS HTG; HTGS_PHASE1.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Muzny D.M., Adams C., Adio-Oduola B., Ali-osman, F.R., Allen C.,
Alsbrooks S.L., Amaralunge H.C., Are J.R., Ayele M., Banks T.,
Bardara J., Benton J., Blinag K., Blankenburg K., Bonin D.,
Bouck J., Bowle S., Brileva M., Brown E., Brown M., Bryant N.P.,
Buhay C., Burch P., Burkett C., Burrell K.L., Byrd N.C.,
Carroll T.F., Carter M., Cavazos S.R., Chacko J., Chavez D.,
Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C.,
Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R.,
DeJana M.L., Davis C., Davy-Carroll L., Dederich D.A.,
DeJana K.R., Delgado O., Denn A.L., Ding Y., Dinh H.H.,
Douthwaite K.J., Draper H., Dugan-Rocha S., Durbin K.J.,
Eamhart C., Edgar D., Edwards C.C., Elhaj C., Escotto M.,
Falls T., Ferraguto D., Flagg N., Ford J., Foster P., Frantz P.,
Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R.,
Gorrell J.H., Guevara W., Gunaratne P., Hale S., Hamilton K.,

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

```

Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., Hernandez, J.,
Hernandez, O., Hodgson, A., Hughes, M., Holloway, C., Hollins, B.,
Homi, F., Howard, S., Huber, J., Huliy, S., Hune, J., Jackson, L.E.,
Jacobson, B., Jia, Y., Johnson, R., Jolivet, S., Joudah, S.,
Karlssohn, E., Kelly, S., Khan, U., King, L., Kovach, J., Kovar, C.,
Kravovic, J., Kuresh, A., Landry, N., Leal, B., Lewis, L.C., Lewis, L.,
Li, J., Li, Z., Lichtarge, O., Liu, C., Liu, J., Liu, W., Louisgeed, H.,
Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J.,
Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E.,
Massey, E., Mawhney, E., McLeod, M.P., Meador, M., Mei, G., Melcher, M.,
Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S.,
Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenkwo, S., Ogih, M., Okwuonu, G.,
Oraguwe, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rivers, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Schier, S., Scott, G., Shen, H., Shooshitari, N., Sisson, I.,
Sodergren, E., Sonalike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wlezyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 16581)
Morley, K.C.
Direct Submission
Submitted (10-AUG-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: HCMU
Center clone name: RP11-90C1
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 17755 bases at least Q40
Consensus quality: 17936 bases at least Q30
Consensus quality: 181287 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 19 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1 2175: contig of 2175 bp in length
2176 2275: gap of unknown length
2276 4405: contig of 2130 bp in length
4405 4505: gap of unknown length
4506 8401: contig of 3866 bp in length
8402 8501: gap of unknown length
8502 12794: contig of 4293 bp in length
12795 12894: gap of unknown length
12895 15194: contig of 2300 bp in length
15195 15294: gap of unknown length
15295 20240: contig of 4946 bp in length
20241 20340: gap of unknown length
20341 24892: contig of 4552 bp in length
24893 24992: gap of unknown length
24993 30199: contig of 5207 bp in length

```

30200 30299: gap of unknown length
30300 33759: contig of 3460 bp in length
33760 33859: gap of unknown length
33860 38788: contig of 4929 bp in length
38789 38888: gap of unknown length
38889 45969: contig of 7081 bp in length
45970 46069: gap of unknown length
46070 55234: contig of 9165 bp in length
55235 55334: gap of unknown length
56750 66750: contig of 11416 bp in length
66751 81272: contig of 14422 bp in length
81273 81372: gap of unknown length
81373 95615: contig of 14243 bp in length
95616 95715: gap of unknown length
109015 109115: contig of 13300 bp in length
109116 125169: contig of 16054 bp in length
125170 125269: gap of unknown length
125270 146580: contig of 21311 bp in length
146581 146581: gap of unknown length
146581 186581: contig of 39901 bp in length.
Location/Qualifiers
1. 186581
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-90C1"

BASE COUNT 58039 a 33763 c 32807 g 60139 t 1833 others
ORIGIN

Query Match 95.0%; Score 19; DB 1; Length 186581;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 TAGTCTTAGTCTTAGT 20
|||||
Db 53280 TAGTCTTAGTCTTAGT 53298

RESULT 6
AC019958/c
LOCUS
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
AC019958
VERSION AC019958.1 GI:6664939
KEYWORDS HTG: HTGS_PHASE2.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 5421)
Adams, M. and Venter, J.C.
Direct Submission
Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
COMMENT This sequence was identified as CDM:10211530 by the submitter. For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1. 5421
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

BASE COUNT 1558 a 1150 c 1146 g 1567 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 5421;
Best Local Similarity 95.0%; Pred. No. 14;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTACTTCTAGTCTTAGT 20
|||||
Db 435 TTACTTCTAGTCTTAGT 416

RESULT 7
AC015356/c
LOCUS
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
AC015356
VERSION AC015356.1 GI:6435979
KEYWORDS HTG: HTGS_PHASE2.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 14921)
Adams, M. and Venter, J.C.
Direct Submission
Submitted (16-NOV-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
COMMENT This sequence was identified as CDM:1021151 by the submitter. For further information on this sequence e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1. 14921
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

BASE COUNT 4439 a 3179 c 2922 g 4381 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 14921;
Best Local Similarity 95.0%; Pred. No. 16;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTACTTCTAGTCTTAGT 20
|||||
Db 12619 TTACTTCTAGTCTTAGT 12600

RESULT 8
AC019964
LOCUS
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered pieces.
AC019964
VERSION AC019964.1 GI:6664933
KEYWORDS HTG: HTGS_PHASE2.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 15707)
Adams, M. and Venter, J.C.
Direct Submission
Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA
COMMENT This sequence was identified as CDM:10211545 by the submitter. For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
1. 15707
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

Query Match
Best Local Similarity 92.0%; Score 18.4; DB 1; Length 52486;
Pred. No. 18;

18770 18869: gap of 100 bp
18870 19591: contig of 722 bp in length
19592 19691: gap of 100 bp
19692 20438: contig of 747 bp in length
20439 20538: gap of 100 bp
20539 21224: contig of 686 bp in length
21225 21324: gap of 100 bp
21325 22053: contig of 735 bp in length
22060 22159: gap of 100 bp
22160 22883: contig of 724 bp in length
22884 22983: gap of 100 bp
22984 23702: contig of 719 bp in length
23703 23802: gap of 100 bp
23803 24523: contig of 721 bp in length
24524 24623: gap of 100 bp
24624 25349: contig of 726 bp in length
25350 25449: gap of 100 bp
25450 26140: contig of 691 bp in length
26141 26240: gap of 100 bp
26241 26985: contig of 745 bp in length
26986 27085: gap of 100 bp
27086 27797: contig of 712 bp in length
27798 27897: gap of 100 bp
27898 28618: contig of 721 bp in length
28619 28718: gap of 100 bp
28719 29447: contig of 729 bp in length
29448 29547: gap of 100 bp
29548 30277: contig of 730 bp in length
30278 30377: gap of 100 bp
30378 31081: contig of 704 bp in length
31082 31181: gap of 100 bp
31182 31918: contig of 737 bp in length
31919 32018: gap of 100 bp
32019 32756: contig of 738 bp in length
32757 32856: gap of 100 bp
32857 33603: contig of 747 bp in length
33604 33703: gap of 100 bp
33704 34414: contig of 711 bp in length
34415 34514: gap of 100 bp
34515 35219: contig of 705 bp in length
35220 35319: gap of 100 bp
35320 35997: contig of 678 bp in length
35998 36097: gap of 100 bp
36098 36801: contig of 704 bp in length
36802 36901: gap of 100 bp
36902 37640: contig of 739 bp in length
37641 37740: gap of 100 bp
37741 38486: contig of 746 bp in length
38487 38586: gap of 100 bp
38587 39331: contig of 745 bp in length
39332 39431: gap of 100 bp
39432 40119: contig of 688 bp in length
40120 40219: gap of 100 bp
40220 40927: contig of 708 bp in length
40928 41027: gap of 100 bp
41028 41741: contig of 714 bp in length
41742 41841: gap of 100 bp
41842 42568: contig of 727 bp in length
42569 42668: gap of 100 bp
42669 43390: contig of 722 bp in length
43391 43490: gap of 100 bp
43491 44216: contig of 726 bp in length
44217 44316: gap of 100 bp
44317 45058: contig of 742 bp in length
45059 45158: gap of 100 bp
45159 45898: contig of 740 bp in length
45899 45998: gap of 100 bp
45999 46710: contig of 712 bp in length
46711 46810: gap of 100 bp
46811 47557: contig of 747 bp in length

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTAGTCTTAGTCTTAGT 20
Db 28551 TTAGTCTTAGTCTTAGT 28532

RESULT 10
AC020246/c
LOCUS
DEFINITION
Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
ACCESSION
AC020246
VERSION
AC020246.1 GI:6664651
KEYWORDS
HTG; HTGS_PHASE2.
SOURCE
Drosophila melanogaster
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 56822)
Adams, M. and Venter, J.C.
Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
This sequence was identified as CDL1021541 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
Location/Qualifiers
1..56822
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

BASE COUNT 17335 a 11120 c 10936 g 17431 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 56822;
Best Local Similarity 95.0%; Pred. No. 18;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTAGTCTTAGTCTTAGT 20
Db 8373 TTAGTCTTAGTCTTAGT 8354

RESULT 11
AC111998/c
LOCUS
DEFINITION
Rattus norvegicus clone CH230-54H8, *** SEQUENCING IN PROGRESS ***,
71 unordered pieces.
ACCESSION
AC111998
VERSION
AC111998.2 GI:21737365
KEYWORDS
HTG; HTGS_PHASE1.
SOURCE
Norway rat.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 138256)
Muzny, D.M., Adams, C., Adio-Oduola, B., All-osman, F.R., Allen, C.,
Alsbrooks, S.L., Amaralunga, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbieri, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carton, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyle, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Doutwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.T.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,

Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabrieli,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Horn,I.F., Howard,S., Huber,J., Huliyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W., Louised,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenko,S., Oguh,M., Okunnu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,D., Perez,L.,
Peters,L., Pickens,R., Plimus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,
Sodergren,E., Sonaite,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczek,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

Unpublished
Direct Submission
2 (bases 1 to 138256)
Worley,K.C.

Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 138256)
Worley,K.C.

Direct Submission
Submitted (13-JUN-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced g1:18701948.

Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: G02S
Center clone name: CH230-54H8

Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 63539 bases at least Q40
Consensus quality: 69255 bases at least Q30
Consensus quality: 73435 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 71 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1037: contig of 1037 bp in length
* 1038 1137: gap of unknown length
* 1138 2308: contig of 1171 bp in length
* 2309 2408: gap of unknown length

2409 3569: contig of 1161 bp in length
* 3570 3669: gap of unknown length
* 3670 4691: contig of 1022 bp in length
* 4692 4791: gap of unknown length
* 4792 5888: contig of 1097 bp in length
* 5889 5988: gap of unknown length
* 5989 7006: contig of 1018 bp in length
* 7007 7106: gap of unknown length
* 7107 8214: contig of 1108 bp in length
* 8215 8314: gap of unknown length
* 8315 9514: contig of 1200 bp in length
* 9515 9614: gap of unknown length
* 9615 10677: contig of 1063 bp in length
* 10678 11880: contig of 1103 bp in length
* 11881 11980: gap of unknown length
* 11981 13176: contig of 1196 bp in length
* 13177 13276: gap of unknown length
* 13277 14439: contig of 1163 bp in length
* 14440 14539: gap of unknown length
* 14540 15677: contig of 1138 bp in length
* 15678 15777: gap of unknown length
* 15778 17258: contig of 1481 bp in length
* 17259 17358: gap of unknown length
* 17359 19001: contig of 1643 bp in length
* 19002 19101: gap of unknown length
* 19102 20135: contig of 1034 bp in length
* 20136 20235: gap of unknown length
* 20236 21365: contig of 1330 bp in length
* 21366 21665: gap of unknown length
* 21666 23148: contig of 1483 bp in length
* 23149 23248: gap of unknown length
* 23249 24746: contig of 1498 bp in length
* 24747 24846: gap of unknown length
* 24847 25905: contig of 1055 bp in length
* 25906 26005: gap of unknown length
* 26006 27419: contig of 1414 bp in length
* 27420 27519: gap of unknown length
* 27520 28527: contig of 1008 bp in length
* 28528 28627: gap of unknown length
* 28628 29768: contig of 1141 bp in length
* 29769 29868: gap of unknown length
* 29869 31525: contig of 1657 bp in length
* 31526 31625: gap of unknown length
* 31626 32975: contig of 1350 bp in length
* 32976 33075: gap of unknown length
* 33076 34670: contig of 1595 bp in length
* 34671 34770: gap of unknown length
* 34771 36362: contig of 1592 bp in length
* 36363 36462: gap of unknown length
* 36463 38206: contig of 1744 bp in length
* 38207 38306: gap of unknown length
* 38307 39464: gap of unknown length
* 39465 39564: gap of unknown length
* 39565 40936: contig of 1372 bp in length
* 40937 41036: gap of unknown length
* 41037 43139: contig of 2103 bp in length
* 43140 43239: gap of unknown length
* 43240 44723: contig of 1484 bp in length
* 44724 44823: gap of unknown length
* 44824 45945: contig of 1121 bp in length
* 45946 46044: gap of unknown length
* 46045 47815: contig of 1771 bp in length
* 47816 47915: gap of unknown length
* 47916 49711: contig of 1796 bp in length
* 49712 49811: gap of unknown length
* 49812 51338: contig of 1427 bp in length
* 51339 51338: gap of unknown length
* 51339 52552: contig of 1214 bp in length
* 52553 52652: gap of unknown length
* 52653 54565: contig of 1913 bp in length
* 54566 55938: contig of 1273 bp in length

```

* 55939 56038: gap of unknown length
* 56039 57191: contig of 1153 bp in length
* 57191 57291: gap of unknown length
* 57291 58953: contig of 1662 bp in length
* 58953 59053: gap of unknown length
* 59053 60514: contig of 1461 bp in length
* 60514 60614: gap of unknown length
* 60614 61983: contig of 1369 bp in length
* 61983 62083: gap of unknown length
* 62083 63489: contig of 1406 bp in length
* 63489 63589: gap of unknown length
* 63589 65566: contig of 1977 bp in length
* 65566 65666: gap of unknown length
* 65666 67994: contig of 2328 bp in length
* 67994 68094: gap of unknown length
* 68094 69382: contig of 1288 bp in length
* 69382 69482: gap of unknown length
* 69482 72060: contig of 2578 bp in length
* 72060 72160: gap of unknown length
* 72160 73569: contig of 1409 bp in length
* 73569 75357: gap of unknown length
* 75357 75457: gap of unknown length
* 75457 76894: contig of 1437 bp in length
* 76894 76994: gap of unknown length
* 76994 78645: contig of 1651 bp in length
* 78645 78746: gap of unknown length
* 78746 80112: contig of 1367 bp in length
* 80112 80212: gap of unknown length
* 80212 82015: contig of 1803 bp in length

```

```

Query Match      92.08; Score 18.4; DB 1; Length 138256;
Best Local Similarity 95.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Oy 1 TTGTTCTTAGTCTTAGT 20
Db 5073 TTACTCTTAGTCTTAGT 5054

```

```

RESULT 12
AC098044
LOCUS
DEFINITION
Rattus norvegicus clone CH230-167M23, *** SEQUENCING IN PROGRESS
***, 59 unordered pieces.
AC098044
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus

```

```

REFERENCE
AUTHORS
1 (bases 1 to 160054)
Wuzy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbier,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouch,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Doutlaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hoques,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL

```

```

COMMENT

```

```

Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lien,C., Liu,J., Liu,M., Louised,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwo,S., Ogih,M., Okwona,G.,
Oragunye,N., Orlando,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojudoan,I., Rolfe,M., Ruiz,S., Severy,G.,
Scherer,S., Scott,G., Shen,H., Shooshtari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 160054)
Worley,K.C.
Direct Submission
Submitted (23-Oct-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 160054)
Worley,K.C.
Direct Submission
Submitted (12-Jul-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On May 29, 2002 this sequence version replaced gi:17964025.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GGD8
Center clone name: CH230-167M23
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 105329 bases at least Q40
Consensus quality: 112592 bases at least Q30
Consensus quality: 118060 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 59 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1 1365: contig of 1365 bp in length
* 1366 1465: gap of unknown length
* 1466 2646: contig of 1181 bp in length
* 2647 2746: gap of unknown length
* 2747 3832: contig of 1086 bp in length
* 3833 3932: gap of unknown length
* 3933 5098: contig of 1166 bp in length
* 5099 5198: gap of unknown length
* 5199 6873: contig of 1675 bp in length
* 6874 6973: gap of unknown length
* 6974 8089: contig of 1116 bp in length

```

8090 8189: gap of unknown length
 * 8190 9467: contig of 1278 bp in length
 * 9468 9567: gap of unknown length
 * 9568 11402: contig of 1835 bp in length
 * 11403 11502: gap of unknown length
 * 11503 12520: contig of 1018 bp in length
 * 12521 12620: gap of unknown length
 * 12621 13988: contig of 1278 bp in length
 * 13989 15056: gap of unknown length
 * 13999 15156: gap of unknown length
 * 15057 16219: contig of 1063 bp in length
 * 15157 16320: gap of unknown length
 * 16220 17573: contig of 1254 bp in length
 * 16320 17574: gap of unknown length
 * 17574 19606: contig of 1933 bp in length
 * 17674 19706: gap of unknown length
 * 19607 21385: contig of 1679 bp in length
 * 19707 21386: gap of unknown length
 * 21386 23103: contig of 1618 bp in length
 * 21486 23203: gap of unknown length
 * 23104 24401: contig of 1198 bp in length
 * 23204 24501: gap of unknown length
 * 24402 26078: contig of 1577 bp in length
 * 24502 26079: gap of unknown length
 * 26079 27359: contig of 1181 bp in length
 * 26179 27360: gap of unknown length
 * 27360 29534: contig of 2075 bp in length
 * 27460 29535: gap of unknown length
 * 29535 31846: contig of 2212 bp in length
 * 29635 31847: gap of unknown length
 * 31847 33652: contig of 1706 bp in length
 * 31947 33753: gap of unknown length
 * 33653 35106: contig of 1354 bp in length
 * 33753 35206: gap of unknown length
 * 35106 37287: contig of 2081 bp in length
 * 35207 37288: gap of unknown length
 * 37288 39465: contig of 2078 bp in length
 * 37388 39466: gap of unknown length
 * 39466 41428: contig of 1863 bp in length
 * 41429 41528: gap of unknown length
 * 41529 44277: contig of 2749 bp in length
 * 44278 44377: gap of unknown length
 * 44378 46469: contig of 2092 bp in length
 * 44378 46570: gap of unknown length
 * 46570 48768: contig of 2199 bp in length
 * 48769 48868: gap of unknown length
 * 48869 50157: contig of 1289 bp in length
 * 50158 50257: gap of unknown length
 * 50258 52556: contig of 2299 bp in length
 * 52557 52656: gap of unknown length
 * 52657 54427: contig of 1771 bp in length
 * 54428 54527: gap of unknown length
 * 54528 56731: contig of 2204 bp in length
 * 56732 56831: gap of unknown length
 * 56832 59637: contig of 2806 bp in length
 * 59638 59737: gap of unknown length
 * 59738 62836: contig of 3099 bp in length
 * 62837 62936: gap of unknown length
 * 62937 65175: contig of 2239 bp in length
 * 65176 65275: gap of unknown length
 * 65276 66950: contig of 1675 bp in length
 * 66951 67050: gap of unknown length
 * 67051 69090: contig of 2040 bp in length
 * 69091 69190: gap of unknown length
 * 69191 72240: contig of 3050 bp in length
 * 72241 72340: gap of unknown length
 * 72341 74931: contig of 2591 bp in length
 * 74932 75031: gap of unknown length
 * 75032 77759: contig of 2628 bp in length
 * 77660 80552: gap of unknown length
 * 77760 80552: contig of 2793 bp in length
 * 80552 80652: gap of unknown length

80653 84775: contig of 4123 bp in length
 * 84776 84875: gap of unknown length
 * 84876 88523: contig of 3648 bp in length
 * 88524 88623: gap of unknown length
 * 88624 91755: contig of 3132 bp in length
 * 91756 91855: gap of unknown length
 * 91856 94622: contig of 2767 bp in length
 * 94623 94722: gap of unknown length
 * 94723 98734: contig of 4012 bp in length
 * 98735 98834: gap of unknown length
 * 98835 103639: contig of 4805 bp in length
 * 103640 103739: gap of unknown length
 * 103740 108154: contig of 4415 bp in length
 * 108155 108254: gap of unknown length
 * 108255 112250: contig of 3996 bp in length
 * 112251 112350: gap of unknown length
 * 112351 115708: contig of 3358 bp in length
 * 115709 115808: gap of unknown length
 * 115809 120163: contig of 4355 bp in length
 * 120164 120263: gap of unknown length
 * 120264 123938: contig of 3675 bp in length
 * 123939 124038: gap of unknown length
 * 124039 128955: contig of 4917 bp in length

Query Match 92.0%; Score 18.4; DB 1; Length 160054;
 Best Local Similarity 95.0%; Pred. No. 20;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTACTGTTCTTACTT 20
 DB 56311 TTAGTCTTACTGTTCTTACTT 56330

RESULT 13
 AC129283/C
 LOCUS
 DEFINITION
 Rattus norvegicus clone CH230-6117, *** SEQUENCING IN PROGRESS ***,
 67 unordered pieces.
 AC129283
 AC129283.1 GI:22000887
 VERSION
 KEYWORDS
 HTG; HTGS_PHASE1.
 SOURCE
 ORGANISM

Rattus norvegicus
 Norway rat.
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE
 AUTHORS

1 (bases 1 to 183218)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alsdrooks,S.L., Amaraltinge,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbara,J., Benton,J., Blinige,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durdin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorell,J.H., Guevara,W., Gunatere,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
 Homsi,F., Howard,S., Huber,J., Huiyk,S., Hume,D., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratochvic,T., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtenarge,O., Lieu,N., Liu,J., Liu,W., Louised,H.,
 Lozada,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Kapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhney,E., Mcleod,M.P., Meador,M., Mel,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,

Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokwenkwo, S., Ogih, M., Okunolu, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L., L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scheer, S., Scott, G., Shen, H., Shooshari, N., Sisson, I.,
Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansley, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Umanil, K., Vasquez, L., Vera, Y., Villalon, D., Vinson, R., Wang, O.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Wallington, S.,
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G., and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 183218)
Moriarty, K.C.

Direct Submission
Submitted (28-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information -----
Center project name: GBWM
Center clone name: CH230-6117
----- Summary Statistics -----
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 127038 bases at least Q40
Consensus quality: 136608 bases at least Q30
Consensus quality: 143417 bases at least Q20

----- NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 67 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1002: contig of 1002 bp in length
1003
1102: gap of unknown length
1103
2859: contig of 1757 bp in length
2860
2959: gap of unknown length
2960
4118: contig of 1159 bp in length
4119
4218: gap of unknown length
4219
5370: contig of 1152 bp in length
5371
5470: gap of unknown length
5471
6980: contig of 1510 bp in length
6981
7080: gap of unknown length
7081
8087: contig of 1007 bp in length
8088
8187: gap of unknown length
8188
9322: contig of 1135 bp in length
9323
9422: gap of unknown length
9423
10530: contig of 1108 bp in length
10531
11677: gap of unknown length
11678
11777: contig of 1047 bp in length
11778
11777: gap of unknown length
11779
13579: contig of 1802 bp in length
13580
13679: gap of unknown length
13680
15247: contig of 1568 bp in length
15248
15347: gap of unknown length
15348
16962: contig of 1615 bp in length
16963
17062: gap of unknown length
17063
18375: contig of 1313 bp in length

18376
18475: gap of unknown length
18476
20601: contig of 2126 bp in length
20602
20701: gap of unknown length
20702
22273: contig of 1572 bp in length
22274
22374: gap of unknown length
22375
24268: contig of 1895 bp in length
24269
24368: gap of unknown length
24369
25603: contig of 1235 bp in length
25604
25703: gap of unknown length
25704
28261: contig of 2558 bp in length
28262
28362: gap of unknown length
28363
29544: contig of 1183 bp in length
29545
29644: gap of unknown length
29645
30747: contig of 1102 bp in length
30748
30846: gap of unknown length
30847
32515: contig of 1669 bp in length
32516
32615: gap of unknown length
32616
34230: contig of 1615 bp in length
34231
34330: gap of unknown length
34331
36380: contig of 2050 bp in length
36381
36480: gap of unknown length
36481
37495: contig of 1015 bp in length
37496
37595: gap of unknown length
37596
40454: contig of 2859 bp in length
40455
40554: gap of unknown length
40555
42500: contig of 1946 bp in length
42501
42600: gap of unknown length
42601
44710: contig of 2110 bp in length
44711
44810: gap of unknown length
44811
44666: contig of 1656 bp in length
44667
46566: gap of unknown length
46567
48625: contig of 2059 bp in length
48626
48725: gap of unknown length
48726
51220: contig of 2495 bp in length
51221
51320: gap of unknown length
51321
53630: contig of 2310 bp in length
53631
53730: gap of unknown length
53731
56025: contig of 2295 bp in length
56026
56125: gap of unknown length
56126
57758: contig of 1633 bp in length
57759
57858: gap of unknown length
57859
59640: contig of 1782 bp in length
59641
59740: gap of unknown length
59741
61958: contig of 2218 bp in length
61959
62058: gap of unknown length
62059
64128: contig of 2070 bp in length
64129
64228: gap of unknown length
64229
66519: contig of 2291 bp in length
66520
66619: gap of unknown length
66620
68461: contig of 1842 bp in length
68462
68561: gap of unknown length
68562
72070: contig of 3509 bp in length
72071
72170: gap of unknown length
72171
74414: contig of 2244 bp in length
74415
74514: gap of unknown length
74515
77258: contig of 2744 bp in length
77259
77358: gap of unknown length
77359
79616: contig of 2258 bp in length
79617
79716: gap of unknown length
79717
83702: contig of 3986 bp in length
83703
83802: gap of unknown length
83803
85968: contig of 2166 bp in length
85969
86068: gap of unknown length
86069
89793: contig of 3725 bp in length
89794
89893: gap of unknown length
89894
92508: contig of 2615 bp in length
92509
92609: gap of unknown length
92610
94912: contig of 2304 bp in length
94913
95012: gap of unknown length
95013
98756: contig of 3744 bp in length
98757
98856: gap of unknown length
98857
102403: contig of 3547 bp in length
102404
102503: gap of unknown length

* 102504 105312: contig of 2809 bp in length
* 105313 105412: gap of unknown length
* 105413 109381: contig of 3969 bp in length
* 109382 109481: gap of unknown length
* 109482 113778: contig of 4297 bp in length
* 113779 113878: gap of unknown length
* 113879 116796: contig of 2918 bp in length
* 116797 116896: gap of unknown length
* 116897 119239: contig of 2343 bp in length
* 119240 119339: gap of unknown length
* 119340 124072: contig of 4733 bp in length
* 124073 124172: gap of unknown length
* 124173 128079: contig of 3907 bp in length
* 128080 128179: gap of unknown length
* 128180 131021: contig of 2842 bp in length
* 131022 131121: gap of unknown length

Query Match 92.0%; Score 18.4; DB 1; Length 183218;
Best Local Similarity 95.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTACTTCTTACTT 20
Db 33225 TTAGTCTTACTTACTTACTT 33206

RESULT 14
AC027745
LOCUS AC027745 192518 bp DNA linear HTG 04-OCT-2000
DEFINITION Homo sapiens clone RP11-535A15, WORKING DRAFT SEQUENCE, 34
unorderd pieces.
ACCESSION AC027745
VERSION AC027745.2 GI:10567968
KEYWORDS HTG, HTGS_PHASE1, HTGS_DRAFT.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 192518)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-535A15
JOURNAL Unpublished
RECORD 2 (bases 1 to 192518)
REFERENCE 2
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,Y., Bede,F.,
Boguski,V., Bouckgealter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hages,B., Heaford,A., Horton,L.,
Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,U., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,
McCarthy,M., McEwan,P., McGuire,A., McKernan,K., McKeeters,R.,
Meldrum,J., Menus,L., Mihova,T., Miranda,C., Mienna,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Olivari,T., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Theodore,J., Titrill,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
TITLE Direct Submission
JOURNAL Submitted (01-APR-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Oct 4, 2000 this sequence version replaced g1:782564.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L8168
Center clone name: 535_A15
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 171261 bases at least Q40
Consensus quality: 181977 bases at least Q30
Consensus quality: 186186 bases at least Q20
Insert size: 210000; agarose-fp
Insert size: 189218; sum-of-contigs
Quality coverage: 3.1 in Q20 bases; agarose-fp
Quality coverage: 3.4 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2775 2874: gap of 100 bp
2775 2874: gap of 100 bp
2875 4195: contig of 1321 bp in length
4196 4295: gap of 100 bp
4296 5880: contig of 1583 bp in length
5881 5980: gap of 100 bp
5981 7755: contig of 1775 bp in length
7756 7855: gap of 100 bp
7856 9023: contig of 1168 bp in length
9024 9123: gap of 100 bp
9124 11085: contig of 1962 bp in length
11086 11185: gap of 100 bp
11186 12742: contig of 1557 bp in length
12743 12842: gap of 100 bp
12843 14362: contig of 1520 bp in length
14363 14462: gap of 100 bp
14463 15891: contig of 1429 bp in length
15892 15991: gap of 100 bp
15992 18140: contig of 2149 bp in length
18141 18240: gap of 100 bp
18241 20252: contig of 2012 bp in length
20253 20352: gap of 100 bp
20353 23974: contig of 3622 bp in length
23975 24074: gap of 100 bp
24075 27028: contig of 2954 bp in length
27029 27128: gap of 100 bp
27129 30517: contig of 3389 bp in length
30518 30617: gap of 100 bp
30618 33387: contig of 2770 bp in length
33388 33487: gap of 100 bp
33488 36476: contig of 2989 bp in length
36477 36576: gap of 100 bp
36577 40395: contig of 3819 bp in length
40396 40495: gap of 100 bp
40496 43638: contig of 3143 bp in length
43639 43738: gap of 100 bp
43739 49735: contig of 5997 bp in length
49736 49835: gap of 100 bp
49836 53882: contig of 4047 bp in length
53883 53982: gap of 100 bp
53983 59666: contig of 5664 bp in length
59667 59766: gap of 100 bp
59767 64782: contig of 5016 bp in length
64783 64882: gap of 100 bp
64883 70725: contig of 5843 bp in length
70726 70825: gap of 100 bp
70826 77243: contig of 6418 bp in length

```

* 77244 77343: gap of 100 bp
* 77344 84177: contig of 6834 bp in length
* 84178 84277: gap of 100 bp
* 84278 91428: contig of 7151 bp in length
* 91429 91528: gap of 100 bp
* 91529 101364: contig of 9836 bp in length
* 101365 101464: gap of 100 bp
* 101465 110674: contig of 9210 bp in length
* 110675 110774: gap of 100 bp
* 110775 120859: contig of 10085 bp in length
* 120860 120959: gap of 100 bp
* 131086 131085: contig of 10126 bp in length
* 131086 131185: gap of 100 bp
* 131186 144973: contig of 13788 bp in length
* 144974 145073: gap of 100 bp
* 145074 158411: contig of 13338 bp in length
* 158412 158511: gap of 100 bp
* 158512 173715: contig of 15204 bp in length
* 173716 173815: gap of 100 bp
* 173816 192518: contig of 18703 bp in length.

FEATURES
source
1..192518
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="RP11-535A15"
/clone_id="RP11 Human Male BAC"
1..2774
/feature="assembly_fragment"
clone_end:SP6
vector_side:left"
2875..4195
/feature="assembly_fragment"
4236..5880
/feature="assembly_fragment"
5981..7755
/feature="assembly_fragment"
7856..9023
/feature="assembly_fragment"
9124..11085
/feature="assembly_fragment"
11186..12742
/feature="assembly_fragment"
12843..14362
/feature="assembly_fragment"
14463..15891
/feature="assembly_fragment"
15992..18140
/feature="assembly_fragment"
18241..20252
/feature="assembly_fragment"
20353..23974
/feature="assembly_fragment"
24075..27028
/feature="assembly_fragment"
27129..30517
/feature="assembly_fragment"
30618..33387
/feature="assembly_fragment"
33488..36476
/feature="assembly_fragment"
36577..40395
/feature="assembly_fragment"
40496..43638
/feature="assembly_fragment"
43739..49735
/feature="assembly_fragment"
49836..53882
/feature="assembly_fragment"
53983..59666
/feature="assembly_fragment"
59767..64782
/feature="assembly_fragment"
64883..70725

```

```

misc_feature /note="assembly_fragment"
70826..77243
misc_feature /note="assembly_fragment"
77344..84177
misc_feature /note="assembly_fragment"
84278..91428
misc_feature /note="assembly_fragment"
91529..101364
misc_feature /note="assembly_fragment"
101365..110674
misc_feature /note="assembly_fragment"
110675..120859
misc_feature /note="assembly_fragment"
120860..131085
misc_feature /note="assembly_fragment"
131086..131185
misc_feature /note="assembly_fragment"
131186..144973
misc_feature /note="assembly_fragment"
144974..145073
misc_feature /note="assembly_fragment"
145074..158411
misc_feature /note="assembly_fragment"
158412..158511
misc_feature /note="assembly_fragment"
158512..173715
misc_feature /note="assembly_fragment"
173716..173815
misc_feature /note="assembly_fragment"
173816..192518

Query Match 92.0% Score 18.4; DB 1; Length 192518;
Best Local Similarity 95.0%; Pred No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTCTTCTTACTTACTT 20
Db 164920 TTAGTCTTCTTACTTACTT 164939

RESULT 15
AC099584/c
LOCUS
DEFINITION Mus musculus clone RP23-1A12, WORKING DRAFT SEQUENCE, 12 unordered
pieces.
ACCESSION AC099584.1 GI:16946033
VERSION AC099584.1
KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT; HTGS-FULLTOP.
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 217395)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguski,K., Bouckgeater,B.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Collangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heathord,A., Horton,L., Hulme,W., Illiev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K.,
Lamazares,R., Landers,T., Lehoczek,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R.,
Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green,P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L17136
Center clone name: 1_A12
----- Summary Statistics

```

Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 214412 bases at least Q40
 Consensus quality: 215767 bases at least Q30
 Consensus quality: 216139 bases at least Q20
 Insert size: 212000; agarose-fp
 Insert size: 216295; sum-of-contigs
 Quality coverage: 9.6 in Q20 bases; agarose-fp
 Quality coverage: 9.4 in Q20 bases; sum-of-contigs

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1      51897: contig of 51897 bp in length
*      51898 51997: gap of 100 bp
*      51998 54376: contig of 2379 bp in length
*      54377 54476: gap of 100 bp
*      54477 58077: contig of 3601 bp in length
*      58078 58177: gap of 100 bp
*      58178 61196: contig of 3019 bp in length
*      61197 61296: gap of 100 bp
*      61297 66269: contig of 4973 bp in length
*      66270 66369: gap of 100 bp
*      66370 74862: contig of 8493 bp in length
*      74863 74962: gap of 100 bp
*      74963 86419: contig of 11457 bp in length
*      86420 86519: gap of 100 bp
*      86520 108007: contig of 21488 bp in length
*      108008 108107: gap of 100 bp
*      108108 132537: contig of 24430 bp in length
*      132538 132637: gap of 100 bp
*      132638 159982: contig of 27345 bp in length
*      159983 160082: gap of 100 bp
*      160083 206764: contig of 46682 bp in length
*      206765 206864: gap of 100 bp
*      206865 217395: contig of 10531 bp in length.

```

FEATURES

source

```

/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="RPCI-23 Female Mouse BAC"
1..51897
/note="assembly-fragment
clone_end:SP6
vector_side:left"
misc_feature
51998..54376
/note="assembly-fragment"
54477..58077
/note="assembly-fragment"
58178..61196
/note="assembly-fragment"
61297..66269
/note="assembly-fragment"
66370..74862
/note="assembly-fragment"
74963..86419
/note="assembly-fragment"
86520..108007
/note="assembly-fragment"
108108..132537
/note="assembly-fragment"
132638..159982
/note="assembly-fragment"
160083..206764
/note="assembly-fragment"
206865..217395

```

```

/note="assembly-fragment
clone_end:r7
vector_side:right"
BASE COUNT      63883 a 44235 c 45035 g 63142 t 1100 others
ORIGIN

```

```

Query Match      92.0%: Score 18.4; DB 1; Length 217395;
Best Local Similarity 95.0%: Pred No. 21;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy      1 TTAGTCTTAGTCTTAGTT 20

```

```

Db 135236 TTAGTCTTAGTCTTAGTT 135217

```

```

Search completed: April 6, 2003, 11:50:25
Job time : 463.5 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 10:34:07 ; Search time 243 Seconds

(without alignments)
1069.066 Million cell updates/sec

Title: US-09-980-265-11

Perfect score: 20

Sequence: 1 ttagcttcttagcttcttagct 20

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1994485 seqs, 6494577260 residues 3988970

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl_NOHTG:*
1: gb_ba:*
2: gb_in:*
3: gb_cm:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: em_ba:*
15: em_fun:*
16: em_hum:*
17: em_in:*
18: em_mu:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_vl:*
29: em_htg_hum:*
30: em_htg_inv:*
31: em_htg_other:*
32: em_htg_mus:*
33: em_htg_pln:*
34: em_htg_rtd:*
35: em_htg_mam:*
36: em_htg_vrt:*
37: em_sy:*
38: em_htgo_hum:*
39: em_htgo_mus:*
40: em_htgo_other:*

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	20	5	AX057377	AX057377 Sequence
2	20	100.0	65	5	AX485115	AX485115 Sequence
3	19	95.0	56820	7	DHMR7C10	ALI21804 Drosophila
4	19	95.0	91268	2	AP004895	Lotus_jap
5	19	95.0	191342	2	AC104140	Drosophila
6	19	95.0	225065	9	AC121835	Mus muscu
7	19	95.0	289090	2	AE003424	Drosophila
8	18.4	92.0	20	5	AX057369	Sequence
9	18.4	92.0	20	5	AX057373	Sequence
10	18.4	92.0	20	5	AX057375	Sequence
11	18.4	92.0	20	5	AX057376	Sequence
12	18.4	92.0	20	5	AX057379	Sequence
13	18.4	92.0	20	5	AX057381	Sequence
14	18.4	92.0	36942	8	AC093782	Sequence
15	18.4	92.0	61372	8	AC092116	Sequence
16	18.4	92.0	71321	7	AP004519	Sequence
17	18.4	92.0	78249	7	AP004529	Sequence
18	18.4	92.0	78574	7	AP004536	Sequence
19	18.4	92.0	87900	7	AP004475	Sequence
20	18.4	92.0	100810	7	AF411804	Sequence
21	18.4	92.0	100810	7	AP004962	Sequence
22	18.4	92.0	101270	7	AP004942	Sequence
23	18.4	92.0	115046	7	AP004907	Sequence
24	18.4	92.0	131995	8	AL139093	Human DNA
25	18.4	92.0	151130	9	AC127432	Sequence
26	18.4	92.0	152492	8	AC079863	Sequence
27	18.4	92.0	152545	2	AC010922	Sequence
28	18.4	92.0	158801	8	AC079118	Sequence
29	18.4	92.0	163464	8	AC007462	Sequence
30	18.4	92.0	164826	2	AC099038	Sequence
31	18.4	92.0	170356	2	AC010558	Sequence
32	18.4	92.0	175867	2	AC007976	Sequence
33	18.4	92.0	179016	2	AC018489	Sequence
34	18.4	92.0	190014	7	AP000836	Sequence
35	18.4	92.0	203831	8	AC093295	Sequence
36	18.4	92.0	244164	2	AE003548	Sequence
37	18.4	92.0	300474	2	AE003503	Sequence
38	18.4	92.0	301783	2	AE003499	Sequence
39	18.4	92.0	324484	2	AE003584	Sequence
40	18	90.0	94052	7	AP004903	Sequence
41	18	90.0	97268	7	AP004536	Sequence
42	18	90.0	97268	7	AP004536	Sequence
43	18	90.0	106023	7	AP004467	Sequence
44	18	90.0	123336	8	AC003075	Sequence
45	18	90.0	131741	7	AP004916	Sequence

ALIGNMENTS

RESULT 1
AX057377
LOCUS AX057377 20 bp DNA linear PAT 17-JAN-2001
DEFINITION Sequence 11 from Patent WO0075304.
ACCESSION AX057377
VERSION AX057377.1 GI:12310118
KEYWORDS
SOURCE
ORGANISM
synthetic construct.
synthetic construct.
artificial sequences.
REFERENCE
AUTHORS Bachy,M., Sodoyer,R. and Tranmy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A II 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
Location/Qualifiers

source 1. .20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"

BASE COUNT 3 a 2 c 3 g 12 t
ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTACTTCTAGTCTTACTT 20
|||||
Db 1 TTACTTCTAGTCTTACTT 20

RESULT 2
AX485115/c AX485115 65 bp DNA linear PAT 16-AUG-2002
DEFINITION Sequence 2415 from Patent WO02053728.
ACCESSION AX485115
VERSION AX485115.1 GI:22319399
KEYWORDS
SOURCE
ORGANISM
Candida albicans.
Candida albicans.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.

REFERENCE
AUTHORS Roemer,T., Jiang,B., Boone,C., Bussey,H. and Ohlsen,K.L.
TITLE Gene disruption methodologies for drug target discovery
JOURNAL Patent: WO 02053728-A 2415 11-JUL-2002;
Elitza Pharmaceuticals, Inc. (US)
FEATURES
source 1. .65
Location/Qualifiers
/organism="Candida albicans"
/db_xref="taxon:5476" 6 t

BASE COUNT 37 a 6 c 16 g 6 t
ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 65;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTACTTCTAGTCTTACTT 20
|||||
Db 46 TTACTTCTAGTCTTACTT 27

RESULT 3
DMB7C10 56820 bp DNA linear INV 07-FEB-2000
LOCUS Drosophila melanogaster BAC clone BACR7C10.
DEFINITION AL121804
ACCESSION AL121804
VERSION AL121804.2 GI:6691820
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster.
Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS Ferraz,C., Vidal,S., Brun,C., Bucheton,A. and Demaille,J.C.
TITLE Sequencing the distal X chromosome of Drosophila melanogaster
JOURNAL Unpublished
REMARK Institut de Genetique Humaine -UPR 1142- CNRS, 141,rue de la
Cardonille, 34396 Montpellier Cedex 5, France
2 (bases 1 to 56820)
REFERENCE
AUTHORS Benos,P.
TITLE Direct Submission
JOURNAL Submitted (06-FEB-2000) European Drosophila Genome Sequencing
Consortium
COMMENT On Jan 12, 2000 this sequence version replaced g1:6018822.
Sequence submitted by Takis Benos, EMBL Outstation - The EBI,

Hinxton, Cambridge, CB10 1SD, U.K.
E-mail: benos@ebi.ac.uk on behalf of the European Drosophila Genome
Sequencing Consortium. For further information see the European
Drosophila Genome Sequencing Consortium's web site:
<http://edgp.ebi.ac.uk/>.
The syntax for the representation of annotation used in this record
is documented at:
ftp://ftp.ebi.ac.uk/pub/databases/edgp/sequence_annotation_README
Coding sequences are predicted from computer analysis, using both
gene and CDS prediction programs and matches to other sequences.
These predictions and matches have been evaluated by the annotators
and may have been refined by hand (in which case a GeneFinder
prediction will have no score. The annotators have also used their
judgement on what matches to represent in this record. A far more
complete annotation record is available from FlyBase
(<http://flybase.bio.indiana.edu/>) through the FlyBase Annotation
Object linked by the db_xref qualifier in the feature table.
IMPORTANT: This sequence is NOT necessarily the entire insert of
clone BACR7C10. It may be shorter, since we are minimising the
overlap between clones to 100 bases, by trimming them. Sequence in
absolute orientation with respect to chromosome clone BACR7C10
overlaps to the left with clone BACR25B3 clone BACR7C10 overlaps to
the right with clone 155E2 the true right end of clone BACR7C10 is
unknown
Clone=BACR7C10; Contig ID=1; Length=56820; Status=Finished.

FEATURES
source 1. .56820
Location/Qualifiers
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/clone="BAC BACR7C10"
join(41..610,672..1487)
/gene="EG:BACR7C10.3"
/note="FBgn0004643:mt(1)15"
join(<41..610,672..1487)
/gene="EG:BACR7C10.3"
/codon_start=1
/protein_id="CAB5854.1"
/db_xref="GI:6691821"
/translation="MRPVADVSYSEADGGTYRLSLSTATKEPSSAHVRPNKQVLN
LRLLHTLAGINCVSVDHVEGIIIDHVKDKMLKLYVECLIPAVESEIEVQSTL
CEDVAOLELVDSFTINPEODRALQDFEYKTYRNNRYRLTARELTIORDLD
MVLAVANNHSAEVAANDPFLPRCMKISADQFKLMDRILOPTDKLGDADPIGV
ISTMLHTYINEVPRVRIKLESTIDQAVLPHNNCMFTTHVAQHNKGIJESLAIAKT
LQATGQOHFRVQVYDOSTILMGTMOEPEFESTITLDSGPKLVKRCRLQDELKNVNA
NVLPEYVNAATFCLELNTFVAELIRRFVLRHSIAQMACLSLDIVLQRAVPLFRE
PNEVVOVLSWIKLQQLKAMLASIMEITELMGVGPLASYSDEIKHIFALFQDT
DMRAKAITQIV"
complement(join(1784..2154,2220..2682,2928..3061,
3129..3370,3429..3625))
/gene="EG:BACR7C10.4"
complement(join(1784..2154,2220..2682,2928..3061,
3129..3370,3429..3625))
/gene="EG:BACR7C10.4"
/note="prediction-method:genefinder",
version:"084", score:"75.04",
/prediction-method:genscan, version:"1.0",
score:"178.97",
/match="desc:2INC-FINGER PROTEIN BOP", species:"Mus
musculus (Mouse)", ranges:(query:3262..3366,
target:SPRMBL:p97443:95..61, score:"109.00"),
(query:2374..2508, target:SPRMBL:p97443:291..247,
score:"98.00"), (query:1844..2146,
target:SPRMBL:p97443:431..331, score:"137.00"),
method:"blastx", version:"1.4.9")"
/codon_start=1
/protein_id="CAB58065.1"
/db_xref="GI:6018870"
/translation="MATRYGSSRSATTSATASCSKSNLKNPAPQIKRGORILTE
KPAFVVKSOYRLRCDNCEATKVLKSCRCRVSYCHRSQQAQAGOHQHEPLIK
VHPRPVADARMLRCRLIRLEHGDLIRGYEHRGKRFDMLSHAEIKNDMPRLIEH
LDSLHVLDMMAESPSTVNPNTKLTGNTFNILADAMNSIAVAILGVST
DHCSCNAVATFEGNELHVAIEDMCLMSKFTFIYIDLANTPEORLDLKEHYEFL
CVSKCTDAKESKEMIALCPNRCAGISVDRNNCPRODAGISPLRNANFNAMLT

```
gene
    RHLNENKDVAYILDVCVKDQGTGVFHPRLNWMYKTLDAFEAEIEVGSKSDALDYG
    QRLPLGRKRYHGPNWNPILGLIHLMKLGRIQLYEGHSKSKALHLLLEAQRILLVTGHRDHR
    LTRFQLYMLVLOAROEAH"
    complement(5405..12456)
    /gene="EG:BACR7C10.5"
    /note="FBgn00101302;Burdock"
    complement(5405..12456)
    /gene="FBgn00101302;Burdock"
    /function="transposon"
    complement(join(13393..13718,13826..13996,14073..14169,
    14175..14489))
    /gene="EG:BACR7C10.6"
    complement(join(13393..13718,13826..13996,14073..14169,
    14175..14489))
    /gene="EG:BACR7C10.6"
    /note="EG:BACR7C10.6"
    /prediction=(method:"genscan", version:"1.0")
    /match=(method:"TETRASPAN NET-4", species:"Homo sapiens
    (Human)", ranges:(query:14274..14375,
    target:SPTRMBL::060746:44..11, score:"96.00"),
    (query:14071..14154, target:SPTRMBL::060746:109..82,
    score:"90.00"), (query:13854..13943,
    target:SPTRMBL::060746:155..126, score:"97.00"),
    (query:13402..13713, target:SPTRMBL::060746:268..165,
    score:"159.00"), method:"blastx", version:"1.4.9"),
    /match=(desc:"CELL SURFACE GLYCOPROTEIN A15 (T-CELL ACUTE
    LYMPHOBLASTIC LEUKEMIA ASSOCIATED ANTIGEN 1) (TALL-1)
    (MEMBRANE COMPONENT, X CHROMOSOME, SURFACE MARKER 1)",
    species:"Homo sapiens (Human)",
    ranges:(query:14289..14375,
    target:SWISS-PROT::P41732:34..6, score:"75.00"),
    (query:14065..14145, target:SWISS-PROT::P41732:103..77,
    score:"53.00"), (query:13845..13928,
    target:SWISS-PROT::P41732:150..123, score:"51.00"),
    (query:13636..13695, target:SWISS-PROT::P41732:175..156,
    score:"56.00"), (query:13444..13548,
    target:SWISS-PROT::P41732:229..195, score:"73.00"),
    method:"blastx", version:"1.4.9")
    /match=(desc:"GH15125.5prime GH Drosophila melanogaster
    head port2 Drosophila melanogaster cDNA clone GH15125
    5prime, mRNA sequence", species:"Drosophila melanogaster
    (fruit fly)", ranges:(query:14153..14694,
    target:EMBL:A1238968:542..1, score:"2710.00"),
    (query:14153..14694, target:EMBL:A138968:542..1,
    score:"2710.00"), (query:14081..14154,
    target:EMBL:A1238968:615..542, score:"370.00"),
    method:"blastn", version:"1.4.9")
    /codon_start=1
    /protein_id="CAB65855.1"
    /db_xref="GI:6691822"
    /translation="MSNRYOGAGIGVGMAAGRGYSQIEVHEWPHHPTVYSCVK
    YMITLNFVFWLFGGLIGVYAFROKMEKANGSVLENTYDFLNLSTVMTLAGTV
    IFTLLRLGALRENTFLKTYSCWLLFLEMAIATVCFPOPMNTEFSCFHKI
    IHSYRDPDQNFIDFAQEFKCCGLSNSGQDMASKNVECSVSKEKCVPPSKCI
    NADISSGLVIMCGYGVONAPVEATKLTWTSCTEIVRWAEHNLVIAGNALGIA
    LIQLVLYLAKTEIGQELQKSRMLA"
    join(15266..15549,15781..16268,16335..16675)
    /gene="EG:BACR7C10.1"
    join(15266..15549,15781..16268,16335..16675)
    /gene="EG:BACR7C10.1"
    /note="EG:BACR7C10.1"
    /prediction=(method:"genscan", version:"1.0",
    score:"171.47")
    /match=(method:"genefinder", version:"084")
    /match=(desc:"LD08769.5prime LD Drosophila melanogaster
    embryo BlueScript Drosophila melanogaster cDNA clone
    LD08769 5prime, mRNA sequence", species:"Drosophila
    melanogaster (fruit fly)", ranges:(query:14795..15334,
    target:EMBL:AA390535:1..540, score:"2691.00"),
    method:"blastn", version:"1.4.9")
    /match=(desc:"SD04409.5prime SD Drosophila melanogaster
    Schneider L2 cell culture port2 Drosophila melanogaster
    cDNA clone SD04409 5prime, mRNA sequence",
    species:"Drosophila melanogaster (fruit fly)",
```

```
ranges:(query:15095..15549, target:EMBL:A1532791:1..455,
score:"2230.00"), (query:15778..15839,
target:EMBL:A1532791:453..514, score:"292.00"),
method:"blastn", version:"1.4.9")
/match=(desc:"HYPOTHEICAL 46.2 KD PROTEIN",
species:"Mus musculus (Mouse)",
ranges:(query:15410..15550,
target:SPTRMBL::0922E9:52..98, score:"139.00"),
(query:15842..16282, target:SPTRMBL::0922E9:123..269,
score:"289.00"), method:"blastx", version:"1.4.9")"
/codon_start=1
/protein_id="CAB65856.1"
/db_xref="GI:6691823"
/translation="MNILRLIFALPDLGGRFLIRPAVNLGMVYDRVRSKADK
VGIVRELRLGLIAFAVILITLAVEMVYAAFYVMPAISHTRPHQKCTLEST
PCTPRAHVSILTKKOQLMVGDAKVIYNIPSPOLIEIGMVCAGKEMDQSMRL
GHSRSAMRRSRPLIRNISTWVSLPLVLCMKKEFOQVPEYFISRYLEQHDITDY
VVEIOSKIQRYFTYLTIVADFTGLRTIMEWVPAISLNLFLFLVFLLSWTH
WSDKMLHSHVQIKYARLRLSLPEGVHSKASLRLDDDLVAYSDDKSIDVGGDTLS
DVIDDDLVLVKKRSRGKRESDALRKPTKRTTADH"
complement(join(17148..18350,18385..18468,18549..19451))
/gene="EG:BACR7C10.7"
complement(join(17148..18350,18385..18468,18549..19451))
/gene="EG:BACR7C10.7"
/note="EG:BACR7C10.7"
/prediction=(method:"genscan", version:"1.0",
score:"356.84")
/match=(method:"genefinder", version:"084")
/match=(desc:"POLYCYSTIC KIDNEY DISEASE AND RECEPTOR FOR
EGG JELLY RELATED PROTEIN", species:"Homo sapiens
(Human)", ranges:(query:18733..18456,
target:SPTRMBL::095850:1928..1901, score:"71.00"),
score:95.0%, Score 19; DB 2; Length 56820;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 19; Conservative 0; Mismatches 0; Gaps 0;
```

QY 1 TTAGTCTTACTTCTTACT 19
DB 34495 TTAGTCTTACTTCTTACT 34513
|||||

RESULT 4
AP004895/C 91268 bp DNA linear PLN 19-JUL-2002
LOCUS Lotusc japonicus genomic DNA, chromosome 4, clone:LJT02D01, TM0039,
DEFINITION complete sequence.
ACCESSION AP004895
VERSION AP004895.1 GI:21907911
KEYWORDS HTG.
SOURCE Lotus japonicus
ORGANISM Lotus japonicus DNA, clone_ltb:LJT library clone:LJT02D01.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Lotaeae;
Lotus.
REFERENCE 1
Kaneko,T., Nakamura,Y., Asamizu,E., Kato,T., Sato,S. and Tabata,S.
Structural Analysis of a Lotus japonicus Genome. I. Sequence
Features and Mapping of Sixty-six TAC clones which cover the 6.7 Mb
Regions of the Genome
Unpublished
2 (bases 1 to 91268)
Nakamura,Y.
DIRECT SUBMISSION
Submitted (26-Mar-2002) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research, 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:yn@kazusa.or.jp,
URL: http://www.kazusa.or.jp, Tel:81-438-52-3935,
Fax:81-438-52-3934)
FEATURES
source Location/Qualifiers
1..91268
/organism="Lotus japonicus"
/db_xref="taxon:34305"

BASE COUNT 27763 a 20258 c 17608 g 25639 t
 ORIGIN /chromosome="4"
 /clone="fj702D01"
 /clone_lib="fj7T_1library"
 /note="TAC clone:TM0039"

Query Match 95.0%; Score 19; DB 7; Length 91268;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 TACTCTTACTCTTACTT 20
 Db 61908 TAGTCTTACTCTTACTT 61890

RESULT 5
 AC104140
 LOCUS 191342 bp DNA linear INV-24-JUL-2002
 DEFINITION Drosophila melanogaster X BAC RP98-9H13 (Roswell Park Cancer
 AC104140
 AC104140.5 GI:17940805
 HTG.
 Drosophila melanogaster.
 ACCESSION
 AC104140
 VERSION
 AC104140.5
 KEYWORDS
 ORGANISM
 Drosophila melanogaster.
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
 AUTHORS
 Muzny,D., Scherer,S., Adams,M.D., Holt,R.A., Evans,C.A.,
 Gocayne,J.D., Taber,P., Williamson,A., Homsli,F.H.,
 Dugan-Rocha,S.D., Sodergren,E.S., Hodgson,A.H., Chen,R.C.,
 Ayele,M.A., Scott,G.S., Worley,K.W., Amandlides,P.G., Brandon,R.C.,
 Rogers,Y., An,H., Baldwin,D., Beeson,K.Y., Brown,M., Buhay,C.,
 Busan,D.A., Center,A., Chen,G., Chen,Z., Clerc-Blankenburg,K.,
 Draper,H., Emery-Cohen,A., Ferrier,S., Garg,N.D.S., Houck,J.,
 Hostin,D., Howland,T.J., Hume,J., Idegawa,C., Jallil,M., Kovar,C.,
 Liu,M., Mattei,B., McIntosh,T.C., Morgan,M., Moy,M., Murphy,B.,
 Nelson,K.A., Ndassa,Y., Nguyen,N., Perez,L., Pittman,G.S., Puri,V.,
 Scheeler,F., Shen,H., Strong,R., Tector,C., Wang,Q., Williams,S.M.,
 Xiang,J., Zaveri,J.S., Zhou,J., Zorrilla,S., Smith,H.O.,
 Wheeler,D., Weinstein,G., Gibbs,R. and Venter,J.C.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 JOURNAL
 Submitted (05-DEC-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 191342)
 Morley,K.C., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alstbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbarella,J.,
 Benton,J., Bimige,K., Blankenburg,K., Bonnin,D., Bouck,J.,
 Bowls,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carton,T.F.,
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
 Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
 Coyle,M.D., Dethorne,S.R., David,R., Davila,M.L., Davis,C.,
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
 Deen,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
 Elhaj,C., Escotto,M., Falls,T., Ferraduto,D., Flagg,N., Ford,J.,
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
 Garza,N., Gill,R., Gorrell,H., Guevara,W., Gunaratne,P., Hale,S.,
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
 Hollins,B., Homsli,F., Howard,S., Huber,J., Huliyk,S., Hume,J.,
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W.,
 Louisedge,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
 Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
 Meier,J., Perez,L., Peters,L., Pickens,R., Plums,E., Pu,L.L.,
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
 Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoosharli,N.,
 Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
 Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellford,B., Thomas,N.,
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
 Wallington,S., Williams,G., Williamson,A., Wleccyk,R., Wooden,S.,
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G., and Gibbs,R.

REFERENCE
 AUTHORS
 JOURNAL
 Submitted (05-DEC-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 191342)
 Morley,K.C., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alstbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbarella,J.,
 Benton,J., Bimige,K., Blankenburg,K., Bonnin,D., Bouck,J.,
 Bowls,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carton,T.F.,
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
 Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
 Coyle,M.D., Dethorne,S.R., David,R., Davila,M.L., Davis,C.,
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
 Deen,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
 Elhaj,C., Escotto,M., Falls,T., Ferraduto,D., Flagg,N., Ford,J.,
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
 Garza,N., Gill,R., Gorrell,H., Guevara,W., Gunaratne,P., Hale,S.,
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
 Hollins,B., Homsli,F., Howard,S., Huber,J., Huliyk,S., Hume,J.,
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W.,
 Louisedge,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
 Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
 Meier,J., Perez,L., Peters,L., Pickens,R., Plums,E., Pu,L.L.,
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
 Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoosharli,N.,
 Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
 Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellford,B., Thomas,N.,
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
 Wallington,S., Williams,G., Williamson,A., Wleccyk,R., Wooden,S.,
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G., and Gibbs,R.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Submitted (20-DEC-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 4 (bases 1 to 191342)
 Morley,K.C., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alstbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbarella,J.,
 Benton,J., Bimige,K., Blankenburg,K., Bonnin,D., Bouck,J.,
 Bowls,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carton,T.F.,
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
 Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
 Coyle,M.D., Dethorne,S.R., David,R., David,R., Davila,M.L., Davis,C.,
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
 Deen,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
 Elhaj,C., Escotto,M., Falls,T., Ferraduto,D., Flagg,N., Ford,J.,
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
 Garza,N., Gill,R., Gorrell,H., Guevara,W., Gunaratne,P., Hale,S.,
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
 Hollins,B., Homsli,F., Howard,S., Huber,J., Huliyk,S., Hume,J.,
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W.,
 Louisedge,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
 Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
 Meier,J., Perez,L., Peters,L., Pickens,R., Plums,E., Pu,L.L.,
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
 Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoosharli,N.,
 Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
 Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellford,B., Thomas,N.,
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
 Wallington,S., Williams,G., Williamson,A., Wleccyk,R., Wooden,S.,
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G., and Gibbs,R.

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 Submitted (27-DEC-2001) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 5 (bases 1 to 191342)
 Morley,K.C., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alstbrooks,S.L., Amaralunge,H.C., Are,J.R., Banks,T., Barbarella,J.,
 Benton,J., Bimige,K., Blankenburg,K., Bonnin,D., Bouck,J.,
 Bowls,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P., Buhay,C.,
 Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carton,T.F.,
 Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R.,
 Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C.,
 Coyle,M.D., Dethorne,S.R., David,R., Davila,M.L., Davis,C.,
 Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O.,
 Deen,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H.,
 Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C.,
 Elhaj,C., Escotto,M., Falls,T., Ferraduto,D., Flagg,N., Ford,J.,
 Foster,P., Frantz,P., Gabisi,A., Gao,J., Garcia,A., Garner,T.,
 Garza,N., Gill,R., Gorrell,H., Guevara,W., Gunaratne,P., Hale,S.,
 Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A.,
 Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C.,
 Hollins,B., Homsli,F., Howard,S., Huber,J., Huliyk,S., Hume,J.,
 Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S.,
 Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,W.,
 Louisedge,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
 Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
 Meier,J., Perez,L., Peters,L., Pickens,R., Plums,E., Pu,L.L.,
 Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M.,
 Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoosharli,N.,
 Sisson,I., Sodergren,E., Sonaike,T., Sparks,A., Stanley,H.,
 Stone,H., Sutton,A., Svatek,A., Taber,P., Tamerisa,A., Tamerisa,K.,
 Tang,H., Tansey,J., Taylor,C., Taylor,T., Tellford,B., Thomas,N.,
 Thomas,S., Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R.,
 Wall,R., Wang,S., Ward-Moore,S., Warren,R., Washington,C.,
 Wallington,S., Williams,G., Williamson,A., Wleccyk,R., Wooden,S.,
 Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G., and Gibbs,R.

Carter, M., Cavazos, S. R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C. D., Cox, C., Coyle, M. D., Dathore, S. R., David, R., Davila, M. L., Davis, C., Davy-Carroll, L., Dederich, D. A., Delaney, K. R., Delgado, O., Dem, A. L., Ding, Y., Din, H. H., Douthwaite, K. J., Draper, H., Dugan-Rocha, S., Durbin, K. J., Earnhart, C., Edgar, D., Edwards, C. C., Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R., Gorrell, J. H., Guevara, M., Gunaratne, P., Hale, S., Hamilton, K., Harris, C., Harris, K., Hart, M., Haylak, P., Hawes, J., Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C., Hollins, B., Homs, F., Howard, S., Huber, J., Huik, S., Hume, J., Jackson, L. E., Jacobson, B., Jia, Y., Johnson, R., Jollivel, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L. C., Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W., Loulseged, H., Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A., Martinez, E., Massey, E., Mawhinney, E., McLeod, M. P., Meador, M., Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K., Morgan, M., Morris, S., Moser, N., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S., Oguh, M., Okunnu, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Picken, R., Primus, E., Pu, L. L., Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojuben, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoshart, N., Sisson, I., Sodergren, E., Sonalke, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, R., Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S., Williams, G., Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y. F., Zhou, J., Zorrilla, S., Nelson, D., Weinstock, G., and Gibbs, R.

Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Dec 20, 2001 this sequence version replaced gi:17736993.
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base

Query Match 95.0%; Score 19; DB 2; Length 191342;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGT 19
Db 143942 TTAGTCTTAGTCTTAGT 143960

RESULT 6
AC121835
LOCUS AC121835 225065 bp DNA linear ROD 13-AUG-2002
DEFINITION Mus musculus chromosome 8 clone RP24-72L18, complete sequence.
ACCESSION AC121835
VERSION AC121835.2 GI:22213523
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE
AUTHORS McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (21-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 225065)
McPherson, J.D. and Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (13-AUG-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
On Aug 13, 2002 this sequence version replaced gi:21039928.

COMMENT

FEATURES
source 1..225065
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="8"
/clone="RP24-72L18"

BASE COUNT 60052 a 50559 c 50972 g 63482 t
ORIGIN

Query Match 95.0%; Score 19; DB 9; Length 225065;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGT 19
Db 148339 TTAGTCTTAGTCTTAGT 148357

RESULT 7
AE003424
LOCUS AE003424 289090 bp DNA linear INV 04-OCT-2000
DEFINITION Drosophila melanogaster genomic scaffold 142000013386054 section 8 of 35, complete sequence.
ACCESSION AE003424 AE002566
VERSION AE003424.1 GI:7290316
KEYWORDS HTG.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera: Endopterygota; Diptera: Brachycera; Muscomorpha; Ephydroidea: Drosophilidae: Drosophila.

1 (bases 1 to 289090)

ADAMS,M.D., CELNIKER,S.E., HOLT,R.A., EVANS,C.A., GOCAYNE,J.D., AMANATIDES,P.G., SCHERER,S.E., LI,P.W., HOSKINS,R.A., GALLE,R.F., GEORGE,R.A., LEWIS,S.E., RICHARDS,S., ASHBURNER,M., HENDERSON,S.N., SUTTON,G.G., WORTMAN,J.R., YANDELL,M.D., ZHANG,Q., CHEN,L.X., BRANDON,R.C., ROGERS,Y.H., BLAZEL,R.G., CHAMPE,M., PEIFFER,B.D., WAN,K.H., DOYLE,C., BAXTER,E.G., HEIT,G., NELSON,C.R., GABOR MIKLOS,G.L., ABRIL,J.F., AGBAYAN,A., AN,H.J., ANDREWS-PFANNKUCH,C., BALDWIN,D., BALLEW,R.M., BASU,A., BAXENDALE,J., BAYRATKATOGLU,T., BEASLEY,E.M., BEESON,K.Y., BENOS,P.V., BERNAN,B.P., BHANDARI,D., BOLSHAKOV,S., BORKOVA,D., BOTCHAN,M.R., BOUCK,J., BROKSTEN,P., BROTLIER,P., BURTLIS,K.C., BUSAM,D.A., BUTLER,H., CADIEU,E., CENTER,A., CHANDRA,I., CHERY,J.M., CRAWLEY,S., DAHLKE,C., DAVENPORT,L.B., DAVIES,P., DE PABLOS,B., DELICHER,A., DENG,Z., MAYS,A.D., DEW,I., DIETZ,S.M., DUDSON,K., DOUP,L.E., DOWNES,M., DUGAN-ROCHA,S., DUNKOV,B.C., DUNN,P., DURBIN,K.J., EVANGELISTA,C.C., FERRAZ,C., FERRIERA,S., FLEISCHMANN,W., FOSLER,C., GABRIELIAN,A.E., GARG,N.S., GELBART,M.W., GLASSER,K., GLODEK,A., GONG,F., GORRELL,J.H., GU,Z., GUAN,P., HARRIS,M., HARRIS,N.L., HARVEY,D., HELMAN,T.J., HERNANDEZ,J.R., HOUCK,J., HOSTIN,D., HOUSTON,K.A., HOWLAND,T.J., WEI,M.H., IBOGWAN,C., JALALI,M., KALUSH,F., KARPEN,G.H., KE-Z., KRAVITZ,S., KULP,D., LAI,Z., LASKO,P., DEL,Y., LEVITSKY,A.A., LI,J., LI,Z., LIANG,Y., LIN,X., LIU,X., MATTEL,B., MCINTOSH,T.C., MCLEOD,M.P., MCPHERSON,D., MERKULOV,G., MILSHINA,N.V., MOBARRY,C., MORRIS,J., MOSHREFI,A., MOUNT,S.M., MOY,M., MURPHY,B., MURPHY,L., MUZY,D.M., NELSON,D.L., NELSON,D.R., NELSON,K.A., NIXON,K., NUSSEKERN,D.R., PACLEB,J.M., PALAZZOLO,M., PITTMAN,G.S., PAN,S., POLLARD,J., PURI,V., REESE,M.G., REINERT,K., REMINGTON,K., SAUNDERS,R.D., SCHEELER,F., SHEN,H., SHUE,B.C., SIDEN-KIAMOS,I., SIMPSON,M., SKUPSKI,M.P., SMITH,T., SPIER,E., SPRADLING,A.C., STAPLETON,M., STRONG,R., SUN,E., SVIRSKAS,R., TECTOR,C., TURNER,R., VENTER,E., WANG,A.H., WANG,X., WANG,Z.Y., WASSARMAN,D.A., WEINSTECK,G.M., WEISSENBAUGH,J., WILLIAMS,S.M., WOODAGE,T., WORTLEY,K.C., WU,D., YANG,S., YAO,Q.A., YE,Q., YEH,R.F., ZAVERT,J.S., ZHAN,M., ZHANG,G., ZHAO,Q., ZHENG,L., ZHENG,X.H., ZHONG,F.N., ZHONG,W., ZHOU,X., ZHU,S., ZHU,X., SMITH,H.O., GIBBS,R.A., MYERS,E.W., RUBIN,G.M. and VENTER,J.C.

The genome sequence of *Drosophila melanogaster*

Science 287 (5461), 2185-2195 (2000)

20196006

JOURNAL

TITLE

MDLINE

2 (bases 1 to 289090)

PUBMED

ADAMS,M.D., CELNIKER,S.E., GIBBS,R.A., RUBIN,G.M. and VENTER,J.C.

REFERENCE

TITLE

AUTHORS

Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

JOURNAL

FEATURES

source

Location/Qualifiers

1. 289090

/organism="Drosophila melanogaster"

/db_xref="taxon:7227"

/chromosome="x"

complement(<9373..>16587)

/gene="EG:BACH48C10.4"

/note="CG14047"

/map="2F5-2F5"

/db_xref="FLYBASE:FBgn0040390"

/db_xref="FLYBASE:FBgn0040390"

complement(join(9373..9459,9519..9720,9784..10373,10559..11005,11151..11813,12630..12711,13406..13521,13582..14921,15552..16167,16573..>16587))

/gene="EG:BACH48C10.4"

/product="CT33606"

/db_xref="FLYBASE:FBan0014047"

/db_xref="FLYBASE:FBgn0040390"

complement(join(9373..9459,9519..9720,9784..10373,10559..11005,11151..11813,12630..12711,13406..13521,13582..14921,15552..16167,16573..16587))

/gene="EG:BACH48C10.4"

/note="EG:BACH48C10.4 gene product"

CDS

/codon_start=1

/protein_id="AAE45777.1"

/db_xref="GI:7290317"

/db_xref="FLYBASE:FBan0014047"

/db_xref="FLYBASE:FBgn0040390"

/translation="MDSIIAAGAGAGVVEADGHEPRNVOYFYTYLTFEFOORTNNS HGVHAGTADTRADRSHGNHGHGVAEDELILNDSDEDERITLYWTRQOLEKT LAMMKQGLDQOQOARLAALARSQENVASADELIFNSDEDERITLYWTRQOLEKT KEINLAKENGGLSAPNFGKRLSGVEALSATSDYSTSESGAGVTSYKADDSVG AMNNSRSAGEGMPVITGSRRLHLDPEVYSJLDPNGSCAGVGGSKPLPPRRFVS VAMOCPPRRKANOQKALVTITETARRSSPTGSGSGSAGTSGSVCAGKAAVACHRCPPEDTHAOLSPDKMEHOTCKLSEPKQTPTTLPEKQDEOLSLMLILAOLA PAATLCOGERIPKGENSTPIIAYVPPDSVLKRTSHVNDNGSCSSNGGCTTSTA TTTTQKQROIINIPEDSCDESPLEEPYRMWSLALRFRTWSSGLVPSDDRMDE ADELDDLDSEKNGHNGSPONEDEEDSARALLVONDLGASVSSGVTVNDEGLAS GATNTRGAVASVMSKFFEEESRAFIYELLRNMAAGQOQOHOQTSTSEDEQDECTS GATSGEEMVCTPTSGGNDODOLILNSENTHSPSTSLNDPDTEEMBELMLA PPMATSTRLGLRRRLREPLFEETESDEEKTQOQSDIDIKGATTNHGHLEDSAGS SDESVPAPPLVPRVPPREDPSAEGQENLAMDQDILSSVVRLLPNIATSM GPRPLTRLLPPTTAIDOPPAVLVASSCELTLDQPSVPGDLAAGSIGTTMTSS RPPPLSPAKAAAPRIGMNAATMTTKTTTMMTPSRAVSTISASDVTSSGGRAS GRPREVPPPPRRRLILQTDIAGAKTWLGGCRRRORGPKNRYVKDVAINTE YQOLGVSQKRTVYINDRDLPAAGTQARLSHFISNLENGKADPESTPAPVAINE SELPGCSASTIDGCARTHCNAPRMWRPKKAGNAAGPQPRSRGGLDLAMRS NDSQNDPTTVNRRPRLSAGLCGSTNSNPNNSSSSSHGSGGLGGLTRAGVEPY GQOOLVVOAGGSGNAGVGGSGSAGNSTCSSRASKGLDEILARRKITYMSOS RSCORVHETSTTTLAMATEMSMDSEKRDSSWPKRAAAALTRTRSTSDMGAG AISTTTTTTTEAEVLEFEVRSKDLINFIRNRSITEASTSSPNSDCAIGQ POSHSQSSLLLOKLPQGOQADQVSEMAQMSPRKNDKPSLRRLMKQYTKRRRTNS VSOIVAG"

complement(<18933..>26541)

/gene="EG:BACH7M4.1"

/note="CG14045"

/map="2F5-2F5"

/db_xref="FLYBASE:FBan0014045"

/db_xref="FLYBASE:FBgn0040387"

complement(join(<18933..20031,20093..20270,20329..21208,21574..22759,24311..24418,25054..25330,25668..26541))

/gene="EG:BACH7M4.1"

/product="CT33604"

/db_xref="FLYBASE:FBan0014045"

/db_xref="FLYBASE:FBgn0040387"

complement(join(18945..20031,20093..20270,20329..21208,21574..22759,24311..24418,25054..25330,25668..26541))

/gene="EG:BACH7M4.1"

/note="EG:BACH7M4.1 gene product"

/codon_start=1

/protein_id="AAE45778.1"

/db_xref="GI:7290318"

/db_xref="FLYBASE:FBan0014045"

/db_xref="FLYBASE:FBgn0040387"

/translation="MPTMTRMHRHSSSAVAVESRGRRGVPGVGDANKENFGVHF MSPFGNASLIALODLSNVHGSPPORSFEGSGPRQATQOLALRLPPTTGGAVAL EDSOLSSRMGDITLMDLALITISARKRCRCTLPGACATGTTSTLIANAMSESS VHEMEVTPPHLAKRQVRYRKRNPKHTTINQROSHAKCLEPDLQVPSRKTSFSSS SASDDDEDEQVAKRSSLASPTTTPPSKCTTTTSSISSSSNGCADMASRGESID VSIAPFAKEQOLNVHVTYRCDLQRSHSGSGSINAIVKVALSGADPGVGHSVGSSE MSSGYOTAVHRHSGRPYDQRRNFOISSEETAGQVQLDLAVHRRHNLNVYQSE MOARVKSGETRTERAERLAKLENAGNEKRSEFLGCSFPLNELVHRHDSVASYKH AOCAPPTSRHSQOPKANAGODQOETEAADODKQODSDPAEAAAVTVYPOKVAIG SGSGNSAANAALNDEVYTSISIDSKGEDLLPQOPOPKLSKKALHQRADNLTPLR FLBDDPADDNANSTTQAOATOSQASKANSNANHLNCTSGGRSTMPNSGGS SVGGAARQOQGRPPMTKRLTTEREGFEFSLVNHPVERTEIGLSADRGILDPG DVIYFDKHNHVMIPADVYLNILRSQSSLTLEIFRSGAGATITSDIGQNNHVS TRGAAVAGSSEHTLATATGTTWMSLORTSTROPANSMSPATVCSGTSSTI EAKRRRLDPOVTFESKESTVYPTVNDNRRLTLISREONMTALHGVDFVPOPLDER KDILSPNDHRTLFOINDELIRLAIEDLLEOLCSSDQOQDEQONMFASRVYLSKTAICA AYKRYCNGIKRACVAVNKSQGTSEFIAITTEPAVPRKPDLTMTIHRLOHREIL KIMQOLAGKPTTEHKNESVTINELQADVREITVSSGMLAGSGLRPLTLQDLES RAVETKCKPPTLAVOGRWFQGLDLSREVERSVKPYWTLLFSDIIFEAKSRPRVFI TEEPTRIANVVDSCFHMRRKTTTEFRLLVDPNGSLASPGVLPDLNLTLPDGLRARKS LTRAPSLTELKAYMQLOROFILVNAALGSPSLSPDSPLSDVLTNLTLPDGLRTTA SMSGSKLPDSLDSIHLNQQOQOQOVRRLRSKRKLDSDGSDYENLQRLASAVRHHCATTTSS

ATGTQSHPOPLCTTSLPRTNSPARNRHPCGHHGVNGVSGFCGCTNNS
CLSSASLITOTQVHQTILNHSOTQSVLSISCKCTVTPETVSEPVNHA
SOKLEFSISVGRSFIDSGGVGVGQSLTTPPTPTSPQPNASFNDFV
ATSTTGKPRNPLAFPGGSMPLDLIDLHEVNDPSTDTVEECIFLGDDEDRGTE
GHHDESDDEDSDVVDPPGMLPPTPPPSLDL"

gene
/db_xref="EG:BACH7M4.4"
/note="CG12496"
/evidence=not_experimental
/map="2f6-3A1"
/db_xref="FLYBASE:FBan0012496"
/db_xref="FLYBASE:FBgn0040385"
/complement(<53568..>54629)
/gene="EG:BACH7M4.4"
/product="CT33334"
/evidence=not_experimental
/db_xref="FLYBASE:FBan0012496"
/db_xref="FLYBASE:FBgn0040385"
/complement(53568..54629)
/gene="EG:BACH7M4.4"
/note="EG:BACH7M4.4 gene product"
/codon_start=1
/evidence=not_experimental
/protein_id="AAF45779.1"
/db_xref="GI:7290319"
/db_xref="FLYBASE:FBan0012496"
/db_xref="FLYBASE:FBgn0040385"
/translation="MIKWTASAOIRLNPVDOANQPORRLORRORRATANK
ENVPDYGYSTPVPITRIRNSPLVLAPLSTIRNTPAVRSQAPQVQSVRTATIL
PREADYSPNGLPFGQHLALRGLAPLDLFIQPRIVSGAGVNLKORLDEDEVAT
LEPLQETAVPEPTPKVEVRSTPQASPTQMGDTDLRLIDALIDSACKNASSKK
PRSTFNLRRRLVQOMESNCPQEVSPSYAPGDDPSDLSEGLVPLPMLNMPTE
PASPLAKIPHLVLQATPVAPKAPCDNTECLEPVALKRTREIYAKKSPIKRYK
VDERNYPEVGLALPSITYV"
complement(<70555..>71976)
/gene="gt"
/note="CG7952"
/map="3A1-3A1"
/db_xref="FLYBASE:FBan0007952"
/db_xref="FLYBASE:FBgn0001150"
complement(join(<70555..71849,71925..>71976))
/gene="gt"
/product="CT23972"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."

Query Match 95.0%; Score 19; DB 2; Length 289090;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTAGTCTTAGTCTTAGT 19
|||||
Db 283054 TTAGTCTTAGTCTTAGT 283072

RESULT 8
AX057369 20 bp DNA linear PAT 17-JAN-2001
LOCUS AX057369
DEFINITION Sequence 3 from Patent WO0075304.
ACCESSION AX057369
VERSION AX057369.1 GI:12310110
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Trannoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 3 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
Location/Qualifiers
1..20
/organism="synthetic construct"

/db_xref="taxon:32630"
/note="Oligonucleotide"

BASE COUNT 4 a 1 c 3 g 12 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTCTTAGTCTTAGT 20
|||||
Db 1 TTAGTATATAGTCTTAGT 20

RESULT 9
AX057373 20 bp DNA linear PAT 17-JAN-2001
LOCUS AX057373
DEFINITION Sequence 7 from Patent WO0075304.
ACCESSION AX057373
VERSION AX057373.1 GI:12310114
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Trannoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 7 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide"

BASE COUNT 3 a 1 c 3 g 13 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTCTTAGTCTTAGT 20
|||||
Db 1 TTAGTCTTAGTCTTAGT 20

RESULT 10
AX057375 20 bp DNA linear PAT 17-JAN-2001
LOCUS AX057375
DEFINITION Sequence 9 from Patent WO0075304.
ACCESSION AX057375
VERSION AX057375.1 GI:12310116
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
artificial sequences.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Trannoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 9 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="Oligonucleotide"

BASE COUNT 4 a 1 c 3 g 12 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGTT 20
Db 1 TTAGTCTTAGTCTTAGTT 20

RESULT 11
AX057376 20 bp DNA linear PAT 17-JAN-2001
LOCUS Sequence 10 from Patent WO0075304.
ACCESSION AX057376
VERSION AX057376.1 GI:12310117
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 10 14-DEC-2000;
Aventis Pasteur (FR)

FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"

BASE COUNT 3 a 1 c 3 g 13 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGTT 20
Db 1 TTAGTCTTAGTCTTAGTT 20

RESULT 12
AX057379 20 bp DNA linear PAT 17-JAN-2001
LOCUS Sequence 13 from Patent WO0075304.
ACCESSION AX057379
VERSION AX057379.1 GI:12310120
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 13 14-DEC-2000;
Aventis Pasteur (FR)

FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"

BASE COUNT 3 a 1 c 4 g 12 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGTT 20
Db 1 TTAGTCTTAGTCTTAGTT 20

RESULT 13
AX057381 20 bp DNA linear PAT 17-JAN-2001
LOCUS Sequence 15 from Patent WO0075304.
DEFINITION

ACCESSION AX057381 GI:12310122
VERSION AX057381.1
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 15 14-DEC-2000;
Aventis Pasteur (FR)

FEATURES
source Location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"

BASE COUNT 3 a 1 c 4 g 12 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 2.7e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGTT 20
Db 1 TTAGTCTTAGTCTTAGTT 20

RESULT 14
AC093782 36942 bp DNA linear PRI 01-MAR-2002
LOCUS Homo sapiens BAC clone RP11-20806 from 4, complete sequence.
ACCESSION AC093782 AC009585
VERSION AC093782.3 GI:15920126
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 36942)
Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 36942)
AUTHORS Abbott,S., Cotton,M. and Doeber,A.
TITLE The sequence of Homo sapiens BAC clone RP11-20806
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 36942)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 36942)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (04-OCT-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 36942)
AUTHORS Waterston,R.
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 4, 2001 this sequence version replaced gi:1579641.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@wustl.wustl.edu
----- Summary Statistics

Center project name: H_NH0208006
Drafting Center: WIBR

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPc1-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenoe, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-4519, 2000 bp overlap.
Actual end of this clone is at base position 36942 of RP11-20806.

Sequence derived from a single plasmid subclone, base position 8833 to 8881.

The sequence of AC009585 has been incorporated into AC093782.

FEATURES

Source

Location/Qualifiers
1. .36942
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-20806"
/clone_id="RPc1-11"
275. .349
/rpt_family="MIR"
568. .903
repeat_region
/rpt_family="MaLR"
1269. .1635
repeat_region
/rpt_family="MaLR"
1830. .2019
repeat_region
/rpt_family="L1"
2633. .2653
repeat_region
/rpt_family="AT_rich"
2714. .3197
repeat_region
/rpt_family="L2"
3209. .3653
repeat_region
/rpt_family="MaLR"
3752. .3857
repeat_region
/rpt_family="L2"
4699. .4741
repeat_region
/rpt_family="MaLR"
4839. .5153
repeat_region
/rpt_family="L2"
5623. .5927
repeat_region
/rpt_family="L2"
6614. .6933
repeat_region
/rpt_family="L2"

repeat_region
/rpt_family="L2"
misc_feature
7875. .9228
/note="CpG island (80C-69.6, o/e=0.80, #CpGs=136)"
8042. .8143
misc_feature
/note="similar to EST AI316961 (NID:940322228) uj24g08.y1"
8064. .8306
misc_feature
/note="similar to EST AI383602 (NID:94196383) tc46c05.x1"
8068. .8339
misc_feature
/note="similar to EST BF308711 (NID:911255965)"
8189. .8339
misc_feature
/note="similar to EST AI316961 (NID:940322228) uj24g08.y1"
8245. .8339
misc_feature
/note="similar to EST AI786568 (NID:95334284) uj17e07.y1"
8352. .8719
misc_feature
/note="similar to EST AM027638 (NID:95886394) wv85g12.x1"
10512. .10628
misc_feature
/note="similar to EST AA913032 (NID:93052424) ol34a06.s1"
10517. .10628
misc_feature
/note="similar to EST AI316961 (NID:940322228) uj24g08.y1"
10517. .10628
misc_feature
/note="similar to EST BF308711 (NID:911255965)"
10517. .10617
misc_feature
/note="similar to EST AI786568 (NID:95334284) uj17e07.y1"
10901. .11008
repeat_region
/rpt_family="ERV1"
10907. .10996
repeat_region
/rpt_family="TC)n"
10935. .11023
repeat_region
/rpt_family="ERV1"
13451. .13470
repeat_region
/rpt_family="TC)n"
16302. .16399
repeat_region
/rpt_family="L1"
16400. .16699
repeat_region
/rpt_family="Alu"
16700. .16940
repeat_region
/rpt_family="L1"
17303. .17650
repeat_region
/rpt_family="Alu"
18082. .18352
repeat_region
/rpt_family="L1"
18614. .18895
misc_feature
/note="similar to EST BF762649 (NID:912110549)"
18942. .19369
misc_feature
/note="similar to EST BG977308 (NID:914380043)"
19505. .19923
misc_feature
/note="similar to EST AI936059 (NID:95674929) wo47g12.x1"
21186. .21476
repeat_region
/rpt_family="Alu"
23528. .23925
misc_feature
/note="similar to EST BF328962 (NID:911299710)"
23983. .24018
repeat_region
/rpt_family="AT_rich"
24479. .24854
misc_feature
/note="similar to EST AI907471 (NID:96497901)"
24804. .24854
misc_feature
/note="similar to EST AI786568 (NID:95334284) uj17e07.y1"
25482. .25781
repeat_region
/rpt_family="Alu"
25847. .25868
repeat_region
/rpt_family="MIR"
26045. .26132
repeat_region
/rpt_family="L2"
26578. .26666
repeat_region
/rpt_family="L2"
26956. .26989
repeat_region
/rpt_family="TC)n"
26990. .27097
repeat_region
/rpt_family="TA)n"
27529. .27954
misc_feature
/note="similar to EST AM027811 (NID:95886567) wv60e01.x1"
28028. .28207
repeat_region

repeat_region /rpt_family="MIR"
28306..28419
/rpt_family="L2"
repeat_region 28838..29136
/rpt_family="Alu"
repeat_region 28997..29018
/rpt_family="Al-rich"
repeat_region 29200..29460
/rpt_family="L2"
misc_feature 29573..29782
/note="similar to EST BF758076 (NID:g12105963)"

Query Match 92.0%; Score 18.4; DB 8; Length 36942;
Best Local Similarity 95.0%; Pred. No. 64;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTCTTAGTT 20
|||||
Db 2953 TTAGTCTTAGTCTTAGTT 2972

RESULT 15
AC092116 61372 bp DNA linear PRI 05-MAR-2002
LOCUS AC092116 Homo sapiens chromosome 16 clone CTD-2133M23, complete sequence.
DEFINITION AC092116
ACCESSION AC092116 GI:19111661
VERSION AC092116.2 GI:19111661
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 61372)
AUTHORS DOE Joint Genome Institute.
TITLE Sequencing of Human Chromosome 16
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 61372)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (22-JUN-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 61372)
AUTHORS DOE Joint Genome Institute.
TITLE Direct Submission
JOURNAL Submitted (05-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Mar 5, 2002 this sequence version replaced gi:14522962.
COMMENT Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
GenBank flat file format but are available as part
of this entry's ASN.1 file.

FEATURES

Source Location/Qualifiers
1..61372
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="CTD-2133M23"

BASE COUNT 19233 a 11657 c 11252 g 19230 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 8; Length 61372;
Best Local Similarity 95.0%; Pred. No. 58;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTCTTAGTT 20
|||||
Db 27299 TTAACTCTAGTCTTAGTT 27318

Search completed: April 6, 2003, 11:56:41
Job time : 418 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:21:32 ; Search time 151 Seconds
(without alignments)
298.278 Million cell updates/sec

Title: US-09-980-265-11

Perfect score: 20

Sequence: 1 ttactcttagtcttagtt 20

Scoring table:

IDENTITY-NUC
Gapop 10.0 , Gapept 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.GeneSeq_101002:*

1: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1980.DAT:*

2: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1981.DAT:*

3: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1982.DAT:*

4: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1983.DAT:*

5: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1984.DAT:*

6: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1985.DAT:*

7: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1986.DAT:*

8: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1987.DAT:*

9: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1988.DAT:*

10: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1989.DAT:*

11: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1990.DAT:*

12: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1991.DAT:*

13: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1992.DAT:*

14: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1993.DAT:*

15: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1994.DAT:*

16: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1995.DAT:*

17: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1996.DAT:*

18: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1997.DAT:*

19: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1998.DAT:*

20: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA1999.DAT:*

21: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2000.DAT:*

22: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2001A.DAT:*

23: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2001B.DAT:*

24: /SIDS2/gcgdata/geneSeq/geneSeqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	100.0	20	22	AA28882
2	19	95.0	20	22	AA28871
3	18.4	92.0	20	22	AA28874
4	18.4	92.0	20	22	AA28878
5	18.4	92.0	20	22	AA28880
6	18.4	92.0	20	22	AA28881
7	18.4	92.0	20	22	AA28883
8	18.4	92.0	20	22	AA28886
9	18.4	92.0	4741	23	ABL27984

C	10	18.4	92.0	13798	23	ABL08500	Drosophila melanog
C	11	18.4	92.0	31034	23	ABL13678	Drosophila melanog
C	12	18	90.0	32183	22	AA199267	Human excretory re
C	13	18	90.0	32183	22	AA163617	Human kidney relat
C	14	18	90.0	32192	22	AA199266	Human excretory re
C	15	18	90.0	32192	22	AA163616	Human kidney relat
	16	16.8	84.0	20	22	AA288872	Immunostimulatory
	17	16.8	84.0	20	22	AA288873	Immunostimulatory
	18	16.8	84.0	20	22	AA288875	Immunostimulatory
	19	16.8	84.0	20	22	AA288876	Immunostimulatory
	20	16.8	84.0	20	22	AA288877	Immunostimulatory
	21	16.8	84.0	20	22	AA288879	Immunostimulatory
	22	16.8	84.0	20	22	AA288885	Immunostimulatory
	23	16.8	84.0	20	22	AA288885	Immunostimulatory
	24	16.8	84.0	20	22	AA288887	Immunostimulatory
	25	16.8	84.0	371	22	AA657516	Human immune/haema
	26	16.8	84.0	377	22	AAK56187	Human immune/haema
	27	16.8	84.0	414	23	ABV47200	Human prostate exp
	28	16.8	84.0	490	23	ABV17406	Human prostate exp
	29	16.8	84.0	572	22	AAK08181	Human brain expres
	30	16.8	84.0	572	22	AAK34060	Human bone marrow
	31	16.8	84.0	572	22	AAI39783	Probe #8469 used t
	32	16.8	84.0	572	24	ABS08840	Human genome-deriv
	33	16.8	84.0	749	24	ABO21728	Oligonucleotide fo
C	34	16.8	84.0	749	24	ABO21729	Oligonucleotide fo
	35	16.8	84.0	1069	21	AAK46432	Arabidopsis thalia
	36	16.8	84.0	1071	21	AAK39993	Arabidopsis thalia
	37	16.8	84.0	1240	21	AAK45636	Arabidopsis thalia
	38	16.8	84.0	1242	21	AAK40226	Arabidopsis thalia
C	39	16.8	84.0	4866	23	ABL04084	Drosophila melanog
C	40	16.8	84.0	4920	18	AAV74337	Staphylococcus aur
	41	16.8	84.0	5453	24	ABL70450	Chemically treated
	42	16.8	84.0	5453	24	AAK56143	Human gene regulat
	43	16.8	84.0	7238	24	ABK31481	Signal transductio
	44	16.8	84.0	7238	24	ABL32448	Human immune syste
	45	16.8	84.0	7353	24	ABL32072	Human immune syste

ALIGNMENTS

RESULT 1	
AA28882	
ID	AA28882 standard; DNA; 20 BP.
AC	AA28882;
DT	09-MAY-2001 (first entry)
DE	Immunostimulatory oligonucleotide #11 as vaccine adjuvant.
KW	Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW	prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW	cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW	phosphorothioate; ss.
OS	Synthetic.
FH	Key
FT	modified_base
FT	Location/Qualifiers
FT	1..20
FT	/*tag= a
FT	/note= "contain phosphorothioate internucleotide linkages"
PN	WO200075304-A1.
PD	14-DEC-2000.
PP	08-JUN-2000; 2000WO-FR01566.
PR	08-JUN-1999; 99FR-0007457.
PR	06-AUG-1999; 99FR-0010378.
XX	

PA (AVET) AVENTIS PASTEUR.
 PI Bachy M, Sodayer R, Tranoy E;
 XX WPI; 2001-041317/05.
 DR
 XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
 PT vaccines for human use, induce lymphocyte proliferation and cytokine
 PT secretion -
 XX
 PS Example 4; Page 13; 30pp; French.
 CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
 CC immunostimulatory oligonucleotide of the invention which contains at
 CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
 CC oligonucleotides do not contain any CG dinucleotides in which C are
 CC unmethylated. The oligonucleotides are used as human immunostimulants
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
 CC They induce proliferation of human lymphocytes, induce secretion of
 CC cytokines, especially interleukin-10 or interferon-gamma and increase
 CC expression of the CD86 activation marker or the CD25 cytokine receptor
 CC on human B lymphocytes.
 CC
 XX
 SO Sequence 20 BP; 3 A; 2 C; 3 G; 12 T; 0 other;
 Query Match 100.0%; Score 20; DB 22; Length 20;
 Best Local Similarity 100.0%; Pred. No. 8.8;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 TTACTTCTAGTCTTACTT 20
 Db 1 TTACTTCTAGTCTTACTT 20
 RESULT 2
 AAF28871
 ID AAF28871 standard; DNA: 20 BP.
 AC AAF28871;
 XX
 DT 09-MAY-2001 (first entry)
 DE Generic immunostimulatory oligonucleotide as vaccine adjuvant.
 XX
 KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
 KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
 KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor; ss.
 XX
 OS Synthetic.
 XX
 PN WO200075304-A1.
 PD 14-DEC-2000.
 XX
 PE 08-JUN-2000; 2000MO-FR01566.
 XX
 PR 08-JUN-1999; 99FR-0007457.
 PR 06-AUG-1999; 99FR-0010378.
 XX
 PA (AVET) AVENTIS PASTEUR.
 PI Bachy M, Sodayer R, Tranoy E;
 XX WPI; 2001-041317/05.
 DR
 XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
 PT vaccines for human use, induce lymphocyte proliferation and cytokine
 PT secretion -
 XX
 PS Claim 8; Page 17; 30pp; French.
 CC This sequence represents a generic example of an immunostimulatory
 CC oligonucleotide of the invention which contains at least one sequence

CC 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The oligonucleotides
 CC do not contain any CG dinucleotides in which C are unmethylated.
 CC The oligonucleotides are used as human immunostimulants and as adjuvants
 CC in therapeutic and prophylactic vaccines for human use. They induce
 CC proliferation of human lymphocytes, induce secretion of cytokines,
 CC especially interleukin-10 or interferon-gamma and increase expression
 CC of the CD86 activation marker or the CD25 cytokine receptor on human B
 CC lymphocytes.
 CC
 XX
 SO Sequence 20 BP; 3 A; 1 C; 3 G; 12 T; 1 other;
 Query Match 95.0%; Score 19; DB 22; Length 20;
 Best Local Similarity 95.0%; Pred. No. 24;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Oy 1 TTACTTCTAGTCTTACTT 20
 Db 1 TTACTTCTAGTCTTACTT 20
 RESULT 3
 AAF28874
 ID AAF28874 standard; DNA: 20 BP.
 AC AAF28874;
 XX
 DT 09-MAY-2001 (first entry)
 DE Immunostimulatory oligonucleotide #3 as vaccine adjuvant.
 XX
 KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
 KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
 KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
 KW phosphorothioate; ss.
 XX
 OS Synthetic.
 XX
 PN WO200075304-A1.
 PD 14-DEC-2000.
 XX
 PE 08-JUN-2000; 2000MO-FR01566.
 XX
 PR 08-JUN-1999; 99FR-0007457.
 PR 06-AUG-1999; 99FR-0010378.
 XX
 PA (AVET) AVENTIS PASTEUR.
 PI Bachy M, Sodayer R, Tranoy E;
 XX WPI; 2001-041317/05.
 DR
 XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
 PT vaccines for human use, induce lymphocyte proliferation and cytokine
 PT secretion -
 XX
 PS Example 4; Page 13; 30pp; French.
 CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
 CC immunostimulatory oligonucleotide of the invention which contains at
 CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
 CC oligonucleotides do not contain any CG dinucleotides in which C are
 CC unmethylated. The oligonucleotides are used as human immunostimulants
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
 CC They induce proliferation of human lymphocytes, induce secretion of
 CC cytokines, especially interleukin-10 or interferon-gamma and increase
 CC expression of the CD86 activation marker or the CD25 cytokine receptor

CC on human B lymphocytes.

XX Sequence 20 BP; 4 A; 1 C; 3 G; 12 T; 0 other:

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 22; Length 20;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTCTTAGTT 20
||||| |||||||||
Db 1 TTAGTTATTAGTTCTTAGTT 20

RESULT 4

AAF28878 ID AAF28878 standard; DNA; 20 BP.

XX AAF28878;

DT 09-MAY-2001 (first entry)

XX Immunostimulatory oligonucleotide #7 as vaccine adjuvant.

KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KM prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.

XX Synthetic.

OS Key Location/Qualifiers

FH modified_base 1..20

FT /tag= a
FT /note= "contain phosphorothioate internucleotide linkages"

FT WO200075304-A1.

XX 14-DEC-2000.

PD 08-JUN-2000; 2000WO-FR01566.

XX 08-JUN-1999; 99FR-0007457.

PR 06-AUG-1999; 99FR-0010378.

XX (AVET) AVENTIS PASTEUR.

XX Bachy M, Sodoyer R, Tranoy E;

PI WPI; 2001-041317/05.

DR New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
XX vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
PS Example 4; Page 13; 30pp; French.

XX Oligonucleotides AAF28872-AAF2887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2RT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.

SO Sequence 20 BP; 3 A; 1 C; 3 G; 13 T; 0 other:

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 22; Length 20;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTCTTAGTT 20
||||| |||||||||
Db 1 TTAGTTATTAGTTCTTAGTT 20

RESULT 5

AAF28880 ID AAF28880 standard; DNA; 20 BP.

XX AAF28880;

DT 09-MAY-2001 (first entry)

XX Immunostimulatory oligonucleotide #9 as vaccine adjuvant.

KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KM prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.

XX Synthetic.

OS Key Location/Qualifiers

FH modified_base 1..20

FT /tag= a
FT /note= "contain phosphorothioate internucleotide linkages"

FT WO200075304-A1.

XX 14-DEC-2000.

PD 08-JUN-2000; 2000WO-FR01566.

XX 08-JUN-1999; 99FR-0007457.

PR 06-AUG-1999; 99FR-0010378.

XX (AVET) AVENTIS PASTEUR.

XX Bachy M, Sodoyer R, Tranoy E;

PI WPI; 2001-041317/05.

DR New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
XX vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
PS Example 4; Page 13; 30pp; French.

XX Oligonucleotides AAF28872-AAF2887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2RT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unmethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially interleukin-10 or interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.

SO Sequence 20 BP; 4 A; 1 C; 3 G; 12 T; 0 other:

Query Match

Best Local Similarity 92.0%; Score 18.4; DB 22; Length 20;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTCTTAGTCTTAGTT 20
||||| |||||||||
Db 1 TTAGTTATTAGTTATTAGTT 20

RESULT 6

ID	AAAF28881	standard; DNA; 20 BP.
XX	AAAF28881;	
AC	AAAF28881;	
DT	09-MAY-2001	(first entry)
XX	Immunostimulatory oligonucleotide #10 as vaccine adjuvant.	
DE	Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;	
KW	prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;	
KM	cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;	
KW	phosphorothioate; ss.	
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	modified_base	1..20
FT	/*tag=	a
FT	/note=	"contain phosphorothioate internucleotide linkages"
XX		
PD	MO200075304-AI.	
XX	14-DEC-2000.	
PF	08-JUN-2000; 2000WO-FR01566.	
XX	08-JUN-1999; 99FR-0007457.	
PR	06-AUG-1999; 99FR-0010378.	
XX	(AVET) AVENTIS PASTEUR.	
PA		
PI	Bachy M, Sodoyer R, Tranoy E;	
XX		
DR	WPI; 2001-041317/05.	
XX		
PT	New immunostimulatory oligonucleotides, useful e.g. as adjuvants in	
PT	vaccines for human use, induce lymphocyte proliferation and cytokine	
PT	secretion -	
XX		
PS	Example 4; Page 13; 30pp; French.	
XX		
CC	Oligonucleotides AAF28872-AAF28887 represent specific examples of an	
CC	immunostimulatory oligonucleotide of the invention which contains at	
CC	least one sequence 5'-TTN1N2T1-3' where N1 and N2 are A, T, C or G. The	
CC	oligonucleotides do not contain any CG dinucleotides in which C are	
CC	unmethylated. The oligonucleotides are used as human immunostimulants	
CC	and as adjuvants in therapeutic and prophylactic vaccines for human use.	
CC	They induce proliferation of human lymphocytes, induce secretion of	
CC	cytokines, especially interleukin-10 or interferon-gamma and increase	
CC	expression of the CD86 activation marker or the CD25 cytokine receptor	
CC	on human B lymphocytes.	
XX		
SC	Sequence 20 BP; 3 A; 1 C; 3 G; 13 T; 0 other;	
XX		
SC	Query Match	92.0%; Score 18.4; DB 22; Length 20;
SC	Best Local Similarity	95.0%; Pred. No. 43;
SC	Matches 19; Conservative 0; Mismatches 1;	Indels 0; Gaps 0;
OY	1 TTAGTTCTTAGTCTTAGTT 20	
DB	1 TTAGTCTTAGTCTTAGTT 20	
XX		
XX	RESULT 7	
XX	AAAF28883	
XX	AAAF28883 standard; DNA; 20 BP.	
AC	AAAF28883;	
XX		
DT	09-MAY-2001	(first entry)
XX		

DE		Immunostimulatory oligonucleotide #12 as vaccine adjuvant.
XX		
KM		Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KM		prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KM		cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KM		phosphorothioate; ss.
OS		Synthetic.
XX		
FH	Key	Location/Qualifiers
FT	modified_base	1..20
FT	/tag= "a	"contain phosphorothioate internucleotide
FT	/note= "	linkages"
XX		
PN	WO200075304-A1.	
PD	14-DEC-2000.	
XX		
PF	08-JUN-2000; 2000WO-FR01566.	
XX		
PR	08-JUN-1999; 99FR-0007457.	
PR	06-AUG-1999; 99FR-0010378.	
XX		
PA	(AVET) AVENTIS PASTEUR.	
XX		
PI	Bachy M, Sodoyer R, Tyanoy E;	
DR	WPI; 2001-041317/05.	
XX		
PT	New immunostimulatory oligonucleotides, useful e.g. as adjuvants in	
PT	vaccines for human use, induce lymphocyte proliferation and cytokine	
PT	secretion -	
XX		
PS	Example 4; Page 13; 30pp; French.	
XX		
CC	Oligonucleotides AAF28872-AAF28887 represent specific examples of an	
CC	immunostimulatory oligonucleotide of the invention which contains at	
CC	least one sequence 5'-TTN1NZT1-3' where N1 and N2 are A, T, C or G. The	
CC	oligonucleotides do not contain any CG dinucleotides in which C are	
CC	unmethylated. The oligonucleotides are used as human immunostimulants	
CC	and as adjuvants in therapeutic and prophylactic vaccines for human use.	
CC	They induce proliferation of human lymphocytes, induce secretion of	
CC	cytokines, especially interleukin-10 or interferon-gamma and increase	
CC	expression of the CD86 activation marker or the CD25 cytokine receptor	
CC	on human B lymphocytes.	
XX		
SQ	Sequence 20 BP; 3 A; 1 C; 4 G; 12 T; 0 other;	
	Query Match	92.0%; Score 18.4; DB 22; Length 20;
	Best Local Similarity	95.0%; Pred. NO. 43;
	Matches 19; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
OY	1 TTAGTTCTTAGTTCTTAGTT 20	
	1 TTAGTTCTTAGTTCTTAGTT 20	
DB		
	RESULT 8	
	.AAF28886	
ID	AAF28886 standard; DNA: 20 BP.	
XX		
AC	AAF28886;	
XX		
DT	09-MAY-2001 (first entry)	
XX		
DE	Immunostimulatory oligonucleotide #15 as vaccine adjuvant.	
KM	Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;	
KM	prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;	
KM	cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;	
KM	phosphorothioate; ss.	
XX		

```
OS Synthetic.
XX
XX Key Location/Qualifiers
FH modified_base 1..20
FT /tag= a
FT /note= "contain phosphorothioate internucleotide
FT linkages"
XX
XX WO200075304-A1.
XX
XX 14-DEC-2000.
XX
XX 08-JUN-2000; 2000WO-FR01566.
XX
XX 08-JUN-1999; 99FR-0007457.
XX 06-AUG-1999; 99FR-0010378.
XX
XX (AVERT ) AVENTIS PASTEUR.
XX
XX Bachy M, Sodayer R, Trannoy E;
XX
XX WPI; 2001-041317/05.
XX
XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
XX vaccines for human use, induce lymphocyte proliferation and cytokine
XX secretion -
XX
XX Example 4; Page 13; 30pp; French.
XX
XX Oligonucleotides AAF28872-AAF28887 represent specific examples of an
XX immunostimulatory oligonucleotide of the invention which contains at
XX least one sequence 5'-TTN1N2TT-3' where N1 and N2 are A, T, C or G. The
XX oligonucleotides do not contain any CG dinucleotides in which C are
XX unmethylated. The oligonucleotides are used as human immunostimulants
XX and as adjuvants in therapeutic and prophylactic vaccines for human use.
XX They induce proliferation of human lymphocytes, induce secretion of
XX cytokines, especially Interleukin-10 or Interferon-gamma and increase
XX expression of the CD86 activation marker or the CD25 cytokine receptor
XX on human B lymphocytes.
XX
XX Sequence 20 BP; 3 A; 1 C; 4 G; 12 T; 0 other;
XX
XX Query Match 92.0%; Score 18.4; DB 22; Length 20;
XX Best Local Similarity 95.0%; Pred. No. 43;
XX Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
OY 1 TTAGTCTTAGTCTTAGTT 20
DB 1 TTAGTCTTAGTCTTAGTT 20

RESULT 9
ABL27984/C
ID ABL27984 standard; DNA; 4741 BP.
XX
XX ABL27984;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 35425.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ds.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 23-MAR-2000; 2000US-191637P.
```

```
PR 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NV.
XX
XX Venter JC, Adams M, Li PMD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 35425; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB57737-AB872072), and the encoded proteins
XX (AB857737-AB872072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 4741 BP; 1155 A; 925 C; 1234 G; 1427 T; 0 other;
XX
XX Query Match 92.0%; Score 18.4; DB 23; Length 4741;
XX Best Local Similarity 95.0%; Pred. No. 50;
XX Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
OY 1 TTAGTCTTAGTCTTAGTT 20
DB 356 TAAGTCTTAGTCTTAGTT 337

RESULT 10
ABL08500/C
ID ABL08500 standard; cDNA; 13798 BP.
XX
XX ABL08500;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 19982.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NV.
XX
XX Venter JC, Adams M, Li PMD, Myers EW;
XX
XX WPI; 2001-656860/75.
XX P-PSDB; ABB64397.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 19982; 21pp + Sequence Listing; English.
XX
XX
```

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB13678-AB13672).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.int/pub/published_pct_sequences.
XX
SQ Sequence 13798 BP; 4299 A; 2926 C; 2679 G; 3894 T; 0 other;
Query Match 92.0%; Score 18.4; DB 23; Length 13798;
Best Local Similarity 95.0%; Pred. No. 51;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTACTTCTTACTTCTTACTT 20
Db 9698 TTACTTCTTACTTCTTACTT 9679
||||| |||||||
RESULT 11
AB13678/c
ID AB13678 standard; CDNA; 31034 BP.
XX
AC ABL13678;
XX
DT 26-MAR-2002 (first entry)
XX
DE *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 35516.
XX
KM *Drosophila*; developmental biology; cell signalling; insecticides;
KW pharmaceutical; gene; ss.
XX
OS *Drosophila melanogaster*.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PE 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
DR P-PSDB; ABB69575.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
PS Claim 1; SEQ ID NO 35516; 21pp + Sequence Listing: English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (AB13678-AB13672).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.int/pub/published_pct_sequences.
XX

SQ Sequence 31034 BP; 9528 A; 5898 C; 5991 G; 9617 T; 0 other;
Query Match 92.0%; Score 18.4; DB 23; Length 31034;
Best Local Similarity 95.0%; Pred. No. 52;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTACTTCTTACTTCTTACTT 20
Db 28137 TTACTTCTTACTTCTTACTT 28118
||||| |||||||
RESULT 12
AA199267/c
ID AA199267 standard; DNA; 32183 BP.
XX
AC AA199267;
XX
DT 07-JAN-2002 (first entry)
XX
DE Human excretory related polynucleotide SEQ ID NO 1031.
XX
XX
KM Human; neurotropic; neuroprotective; cytosolic; dermatological; virologic;
KW immunosuppressive; anti-infectious; anti-HIV; antibacterial; vulnereary;
KW antiparkinsonian; antispasmodic; antianemic; antiarthritic; cancer;
KW antitumoral; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrologic; gene therapy; vaccine;
KW excretory system; ds.
XX
OS Homo sapiens.
XX
PN WO200155313-A2.
XX
PD 02-AUG-2001.
XX
PE 17-JAN-2001; 2001WO-US01323.
XX
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 14-AUG-2000; 2000US-0226279.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.

[illegible][illegible]

AA163617/c
ID AA163617 standard; DNA: 32183 BP.
XX
AC AA163617;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human kidney related polynucleotide SEQ ID NO 932.
XX
KW Human: kidney antigen; immunosuppressive; antirheptic; antirheumatic;
KW antiproliferative; cytosolic; cardant; vasotropic; cerebroprotective;
KW neurotropic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; antileptic; hepatotropic; antidiabetic;
KW antiinflammatory; antileptic; vulnery; anticonvulsant; antiparasitic;
KW gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; ds.
XX
OS Homo sapiens.
XX
PN W020015323-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01343.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205415.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216860.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 23-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.

PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 14-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241121.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.

PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0244223.
PR 21-SEP-2000; 2000US-0244274.
PR 25-SEP-2000; 2000US-0244997.
PR 25-SEP-2000; 2000US-0244998.
PR 26-SEP-2000; 2000US-0255464.
PR 27-SEP-2000; 2000US-0255834.
PR 27-SEP-2000; 2000US-0255836.
PR 29-SEP-2000; 2000US-0263637.
PR 29-SEP-2000; 2000US-0263638.
PR 29-SEP-2000; 2000US-0263659.
PR 29-SEP-2000; 2000US-026370.
PR 02-OCT-2000; 2000US-0265802.
PR 02-OCT-2000; 2000US-0270337.
PR 02-OCT-2000; 2000US-0270338.
PR 02-OCT-2000; 2000US-0270339.
PR 02-OCT-2000; 2000US-027040.
PR 13-OCT-2000; 2000US-0293935.
PR 13-OCT-2000; 2000US-0293937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.

PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX WPI, 2001-465569/50.
XX
XX Isolated nucleic acid molecule encoding excretory system antigen is
PT used in preventing, treating or ameliorating a medical condition
XX
XX Example 2; SEQ ID NO 1030; 574pp + Sequence Listing; English.
XX
CC The invention relates to novel excretory system related human
CC polynucleotides (AA198567-AA199503) and the encoded proteins
CC (AAM99594-AAM99913) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy, especially
CC disorders related to the excretory system. The genes are isolated
CC from a range of human tissues disclosed in the specification. The
CC nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIFO at ftp.wipo.int/pub/published_pcl-sequences.
XX
XX Sequence 32192 BP; 8946 A; 6809 C; 6885 G; 9552 T; 0 other;
SQ
Query Match 90.0%; Score 18; DB 22; Length 32192;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TTAGTCTTAGTCTTAG 18
DB 24031 TTAGTCTTAGTCTTAG 24014
RESULT 15
AA163616/c
ID AA163616 standard; DNA; 32192 BP.
XX
XX AA163616;
AC
XX
XX 22-OCT-2001 (first entry)
DT
XX
XX Human kidney related polynucleotide SEQ ID NO 931.
DE
XX
XX Human; kidney antigen; immunosuppressive; antiarthritis; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW nootropic; neuroprotective; antibacterial; virucide; fungicide;
KW ophthalmological; antiallergic; hepatotropic; antidiabetic;
KW antinflammatory; antituber; vulnerable; anticonvulsant; antiparasitic;
KW gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; ds.
KW
XX
XX Homo sapiens.
OS
XX
XX WO200155323-A2.
PN
XX
XX 02-AUG-2001.
PD
XX

PF 17-JAN-2001; 2001WO-US01343.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226868.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 03-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.

PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239335.
PR 13-OCT-2000; 2000US-0239337.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0248207.
PR 17-NOV-2000; 2000US-0248208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.
PR 17-NOV-2000; 2000US-0249214.
PR 17-NOV-2000; 2000US-0249215.
PR 17-NOV-2000; 2000US-0249216.
PR 17-NOV-2000; 2000US-0249217.
PR 17-NOV-2000; 2000US-0249218.
PR 17-NOV-2000; 2000US-0249244.
PR 17-NOV-2000; 2000US-0249245.
PR 17-NOV-2000; 2000US-0249264.
PR 17-NOV-2000; 2000US-0249265.
PR 17-NOV-2000; 2000US-0249297.
PR 17-NOV-2000; 2000US-0249299.
PR 17-NOV-2000; 2000US-0249300.
PR 01-DEC-2000; 2000US-0250160.
PR 01-DEC-2000; 2000US-0250391.
PR 05-DEC-2000; 2000US-0251030.
PR 05-DEC-2000; 2000US-0251988.
PR 05-DEC-2000; 2000US-0256719.
PR 06-DEC-2000; 2000US-0251479.
PR 06-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251856.
PR 08-DEC-2000; 2000US-0251868.
PR 08-DEC-2000; 2000US-0251869.
PR 08-DEC-2000; 2000US-0251989.
PR 08-DEC-2000; 2000US-0251990.
PR 11-DEC-2000; 2000US-0254097.
PR 05-JAN-2001; 2001US-0259678.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX WPI: 2001-488784/53.
XX
PT New isolated nucleic acids and polypeptides, useful for diagnosing,

PT treating and/or preventing human diseases and disorders -
 XX
 PS Disclosure; SEQ ID NO 931; 564pp + Sequence Listing; English.
 XX
 CC The invention relates to novel kidney related polynucleotides
 CC (AA162971-AA163793) and the encoded polypeptides (AA162971-AA163793)
 CC collectively known as kidney antigens and the use of such kidney antigens
 CC for detecting disorders of the kidney, especially kidney cancer and
 CC kidney cancer metastases. The polynucleotides and proteins are also
 CC useful for preventing, treating or ameliorating medical conditions
 CC e.g. by protein or gene therapy. The genes are isolated from a range
 CC of human tissues disclosed in the specification. The nucleic acids,
 CC proteins, antipodes and (ant)agonists are useful in the diagnosis,
 CC treatment and prevention of: (a) cancer, e.g. breast and ovarian cancer,
 CC and other cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital; (b) immune disorders
 CC e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
 CC sclerosis, rheumatoid arthritis and ulcerative colitis;
 CC (c) cardiovascular disorders such as myocardial ischaemia; (d) wound
 CC healing; (e) neurological diseases e.g. cerebral anoxia and epilepsy;
 CC and (f) infectious diseases such as viral, bacterial, fungal and
 CC parasitic infections.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 32192 BP; 8946 A; 6809 C; 6885 G; 9552 T; 0 other;
 Query Match 90.0%; Score 18; DB 22; Length 32192;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 1 TTAGTCTTAGTCTTAG 18
 ||||||||||||||||
 Db 24031 TTAGTCTTAGTCTTAG 24014

Search completed: April 6, 2003, 12:05:51
 Job time : 176 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:12:37 ; Search time 45.25 Seconds
(without alignments)
135.548 Million cell updates/sec

Title: US-09-980-265-11

Perfect score: 20

Sequence: 1 ttagcttctagcttctagct 20

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/lna/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/lna/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/lna/6A.COMB.seq:*
4: /cgn2_6/ptodata/2/lna/6B.COMB.seq:*
5: /cgn2_6/ptodata/2/lna/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/2/lna/Dackfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.4	82.0	111282	4	US-09-754-250-3
2	15.8	79.0	1920	4	US-08-682-767-22
3	15.4	77.0	3162	2	US-08-449-645A-12
4	15.4	77.0	3162	2	US-08-702-367A-12
5	15.4	77.0	3162	5	PCR-US95-04681-12
6	15.2	76.0	995	4	US-08-976-259-101
7	15.2	76.0	1602	2	US-08-770-544-3
8	15.2	76.0	3231	1	US-08-195-152-1
9	15.2	76.0	5000	4	US-09-104-070-3
10	15.2	76.0	20165	4	US-09-609-816-7
11	15.2	76.0	59065	4	US-09-813-817-3
12	15.2	76.0	59065	4	US-09-978-197-3
13	15.2	75.0	1770	3	US-09-006-636-1
14	15.2	75.0	1770	4	US-09-006-636-1
15	15.2	75.0	1770	4	US-09-325-274-1
16	15.2	75.0	6000	3	US-09-006-636-9
17	15.2	75.0	6000	4	US-09-006-636-9
18	15.2	75.0	6000	4	US-09-325-274-9
19	14.8	74.0	524	4	US-09-222-575-169
20	14.8	74.0	98844	4	US-09-791-211-10
21	14.4	72.0	815	3	US-08-433-263B-1
22	14.4	72.0	2011	4	US-09-182-450-2
23	14.4	72.0	2181	4	US-09-780-049-17
24	14.4	72.0	2966	4	US-09-780-049-3
25	14.4	72.0	3381	3	US-08-937-195-1
26	14.4	72.0	3381	3	US-08-937-195-2
27	14.4	72.0	3381	4	US-08-915-152-1

C 28	14.4	72.0	3381	4	US-08-915-152-2	Sequence 2, Appl
C 29	14.4	72.0	3381	5	PCR-US96-07627-1	Sequence 1, Appl
C 30	14.4	72.0	3381	5	PCR-US96-07627-2	Sequence 2, Appl
C 31	14.4	72.0	40000	4	US-09-780-049-18	Sequence 18, Appl
C 32	14.2	71.0	228	1	US-08-222-177A-43	Sequence 43, Appl
C 33	14.2	71.0	362	4	US-09-605-785-442	Sequence 442, App
C 34	14.2	71.0	362	4	US-09-439-313-442	Sequence 442, App
C 35	14.2	71.0	362	4	US-09-352-616A-442	Sequence 442, App
C 36	14.2	71.0	427	4	US-09-556-877-46	Sequence 46, Appl
C 37	14.2	71.0	427	4	US-09-620-412C-46	Sequence 46, Appl
C 38	14.2	71.0	586	4	US-09-385-982-472	Sequence 472, App
C 39	14.2	71.0	1001	4	US-09-641-638-92	Sequence 92, Appl
C 40	14.2	71.0	1001	4	US-09-641-638-93	Sequence 93, Appl
C 41	14.2	71.0	1001	4	US-09-641-638-94	Sequence 94, Appl
C 42	14.2	71.0	1872	4	US-09-280-116-110	Sequence 110, App
C 43	14.2	71.0	1922	1	US-08-700-749A-7	Sequence 7, Appl
C 44	14.2	71.0	1922	3	US-09-020-684-7	Sequence 7, Appl
C 45	14.2	71.0	1922	3	US-09-020-467-7	Sequence 7, Appl

ALIGNMENTS

```

RESULT 1
US-09-754-250-3
; Sequence 3, Application US/09754250
; Patent No. 6376225
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: C1001063
; CURRENT APPLICATION NUMBER: US/09/754,250
; CURRENT FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 111282
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(111282)
; OTHER INFORMATION: n = A,T,C or G
US-09-754-250-3

Query Match      82.0%; Score 16.4; DB 4; Length 111282;
Best Local Similarity 94.4%; Pred. No. 49;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY      3 AGTCTTAGTCTTAGT 20
        |||
Db 102136 AGTCTTAGTCTTAGAT 102153

RESULT 2
US-08-682-767-22
; Sequence 22, Application US/08682767
; Patent No. 6291741
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF MODIFIED PLANTS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of The Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/682,767
FILING DATE: 30-July-96
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: White, John P
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 51735
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0526
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 1920 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
US-08-682-767-22

Query Match 79.0%; Score 15.8; DB 4; Length 1920;
Best local Similarity 89.5%; Pred. No. 76;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAGTCTTAGTCTTAGT 20
Db 48 TATTCTTAGTCTTAGT 66

RESULT 3
US-08-449-645A-12
Sequence 12, Application US/08449645A
Patent No. 5981245
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,645A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2976

US-08-449-645A-12
Query Match 77.0%; Score 15.4; DB 2; Length 3162;
Best local Similarity 94.1%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AGTCTTAGTCTTAGT 19
Db 3144 AGTCTTAGTCTTAGT 3160

RESULT 4
US-08-702-367A-12
Sequence 12, Application US/08702367A
Patent No. 5981246
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California
COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/702,367A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2976
US-08-702-367A-12

Query Match 77.0%; Score 15.4; DB 2; Length 3162;
Best local Similarity 94.1%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AGTCTTAGTCTTAGT 19
Db 3144 AGTCTTAGTCTTAGT 3160

RESULT 5
PCT-US95-04681-12
Sequence 12, Application PC/TUS9504681
GENERAL INFORMATION:
APPLICANT: Fox, Gary M.
TITLE OF INVENTION: EPH-Like Receptor Protein Tyrosine
TITLE OF INVENTION: Kinases
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Amgen Patent Operations/RBW
STREET: 1840 Dehavilland Drive
CITY: Thousand Oaks
STATE: California

COUNTRY: USA
ZIP: 91320
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/04681
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Winter, Robert B.
REFERENCE/DOCKET NUMBER: A-287
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 3162 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2976
PCT-US95-04681-12

Query Match 77.0% Score 15.4; DB 5; Length 3162;
Best Local Similarity 94.1%; Pred. No. 1.1e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 ACTCTTACTTCTTACT 19
Db 3144 ACTCTTACTTCTTACT 3160

RESULT 6
US-08-976-259-101
Sequence 101, Application US/08976259
Patent No. 6316609
GENERAL INFORMATION:
APPLICANT: Dillon, Patrick J.
APPLICANT: Choi, Gail H.
APPLICANT: Welch, Rodney A.
TITLE OF INVENTION: Nucleotide Sequence of Escherichia coli
Patent No. 6316609
NUMBER OF SEQUENCES: 142
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Ave, N.W., Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/976, 259
FILING DATE: Herewith
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,626 AND US 60/061,953
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1488 .0740002/EKS/CM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:

LENGTH: 995 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-976-259-101

Query Match 76.0% Score 15.2; DB 4; Length 995;
Best Local Similarity 85.0%; Pred. No. 1.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGT 20
Db 489 TTAGTCTTAGTCTTAGT 508

RESULT 7
US-08-770-544-3
Sequence 3, Application US/08770544
Patent No. 5907085
GENERAL INFORMATION:
APPLICANT: Gonsalves, Dennis
APPLICANT: Ling, Kai-Shu
TITLE OF INVENTION: GRAPEVINE LEAFROLL VIRUS PROTEINS AND
THEIR USES
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/770,544
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60009008
FILING DATE: 21-DEC-1995
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael L.
REGISTRATION NUMBER: 30,727
REFERENCE/DOCKET NUMBER: 19603/621
TELECOMMUNICATION INFORMATION:
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1602 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-770-544-3

Query Match 76.0% Score 15.2; DB 2; Length 1602;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGT 20
Db 1190 TTAGTCTTAGTCTTAGT 1209

RESULT 8
US-08-195-152-1
Sequence 1, Application US/08195152
Patent No. 5679541

GENERAL INFORMATION:
APPLICANT: Bonini, Nancy M.
APPLICANT: Leiserson, William M.
APPLICANT: Benzer, Seymour
TITLE OF INVENTION: PROGRAMMED CELL DEATH ANTAGONIST
TITLE OF INVENTION: PROTEINS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
STREET: 4 Embarcadero Center, Suite 3400
City: San Francisco
STATE: California
COUNTRY: United States
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/195,152
FILING DATE: 14-FEB-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Trecartin, Richard F.
REGISTRATION NUMBER: 31,801
REFERENCE/DOCKET NUMBER: A-59551/RT/RMS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277289
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3231 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
US-08-195-152-1

Query Match 76.0%; Score 15.2; DB 1; Length 3231;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGT 20
DB 3053 TTAGTTAAAGTCTTAGT 3072

RESULT 9
US-09-104-070-3
Sequence 3, Application US/09104070
Patent No. 6153741
GENERAL INFORMATION:
APPLICANT: Richards, Eric J.
APPLICANT: Jeddell, Jeffrey A.
TITLE OF INVENTION: DNA Methylation Gene from Plants
FILE REFERENCE: Wash U CI-0014 US
CURRENT APPLICATION NUMBER: US/09/104,070
CURRENT FILING DATE: 1998-06-24
EARLIER APPLICATION NUMBER: US 60/083,612
EARLIER FILING DATE: 1998-04-30
NUMBER OF SEQ ID NOS: 3
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 5000
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-104-070-3

Query Match 76.0%; Score 15.2; DB 3; Length 5000;
Best Local Similarity 85.0%; Pred. No. 1.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGT 20
DB 1914 TAAATTCCTGCTTAGT 1933

RESULT 10
US-09-609-816-7/c
Sequence 7, Application US/09609816
Patent No. 6436684
GENERAL INFORMATION:
APPLICANT: Woodage, Trevor
APPLICANT: Wei, Minh Hui
APPLICANT: Kodira, Chinappa
APPLICANT: Beasley, Ellen
APPLICANT: DiFrancesco, Valentina
TITLE OF INVENTION: ISOLATED HUMAN DRUG-METABOLIZING
TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
FILE REFERENCE: CL000669PCT
CURRENT APPLICATION NUMBER: US/09/609,816
CURRENT FILING DATE: 2000-07-03
PRIOR APPLICATION NUMBER: 60/192,408
PRIOR FILING DATE: 2000-03-27
PRIOR APPLICATION NUMBER: 60/212,725
PRIOR FILING DATE: 2000-06-20
PRIOR APPLICATION NUMBER: 09/609,816
PRIOR FILING DATE: 2000-07-03
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 20165
TYPE: DNA
ORGANISM: HUMAN
US-09-609-816-7

Query Match 76.0%; Score 15.2; DB 4; Length 20165;
Best Local Similarity 85.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGT 20
DB 4051 TTGTTTCTGCTTAGT 4032

RESULT 11
US-09-813-817-3
Sequence 3, Application US/09813817
Patent No. 6340583
GENERAL INFORMATION:
APPLICANT: YAN, Chunhua et al.
TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
FILE REFERENCE: CL001178
CURRENT APPLICATION NUMBER: US/09/813,817
CURRENT FILING DATE: 2001-03-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 59065
TYPE: DNA
ORGANISM: Human
US-09-813-817-3

Query Match 76.0%; Score 15.2; DB 4; Length 59065;
Best Local Similarity 85.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGT 20
DB 15314 TTGTTCTTGTATTAGT 15333

RESULT 12
US-09-978-197-3
; Sequence 3, Application US/09978197
; Patent No. 6403353
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: C1001178D1V
; CURRENT APPLICATION NUMBER: US/09/978,197
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 09/813,817
; PRIOR FILING DATE: 2001-03-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 59065
; TYPE: DNA
; ORGANISM: Human
US-09-978-197-3
Query Match 76.0%; Score 15.2; DB 4; Length 59065;
Best Local Similarity 85.0%; Pred. No. 1.5e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 TTAGTCTTACTTCTTACTT 20
DB 15314 TTGTGTTCTTGTATTACTT 15333
- RESULT 13
US-09-006-636-1/c
; Sequence 1, Application US/09006636
; Patent No. 6005092
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Shani, Ziv
; TITLE OF INVENTION: ARABIDOPSIS THALIANA ENDO-1,4-BETA-
; TITLE OF INVENTION: GLUCANASE GENE, PROMOTER AND PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,636
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Baldwin, Geraldine F.
; REGISTRATION NUMBER: 31,232
; REFERENCE/DOCKET NUMBER: 7809-019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1770 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-09-006-636-1

Query Match 75.0%; Score 15; DB 3; Length 1770;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 AGTCTTACTTCTTA 17
DB 659 AGTCTTACTTCTTA 645
- RESULT 14
US-09-006-632-1/c
; Sequence 1, Application US/09006632
; Patent No. 6184440
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Shani, Ziv
; APPLICANT: Shigel, Etai
; TITLE OF INVENTION: TRANSGENIC PLANTS OF ALTERED
; TITLE OF INVENTION: MORPHOLOGY
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/006,632
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Baldwin, Geraldine F.
; REGISTRATION NUMBER: 31,232
; REFERENCE/DOCKET NUMBER: 7809-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1770 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-09-006-632-1
Query Match 75.0%; Score 15; DB 4; Length 1770;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 AGTCTTACTTCTTA 17
DB 659 AGTCTTACTTCTTA 645
- RESULT 15
US-09-325-274-1/c
; Sequence 1, Application US/09325274
; Patent No. 6333023
; GENERAL INFORMATION:
; APPLICANT: Shoseyov, Oded
; APPLICANT: Shani, Ziv
; TITLE OF INVENTION: ARABIDOPSIS THALIANA ENDO-1,4-BETA-
; TITLE OF INVENTION: GLUCANASE GENE, PROMOTER AND PROTEIN
; NUMBER OF SEQUENCES: 21

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSO Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/325,274
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/006,636
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Baldwin, Geraldine F.
; REGISTRATION NUMBER: 31,232
; REFERENCE/DOCKET NUMBER: 7809-019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1770 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
; US-09-325-274-1

Query Match 75.0%; Score 15; DB 4; Length 1770;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 AGTCTTAGTCTTA 17
|||||
Db 659 AGTCTTAGTCTTA 645

Search completed: April 6, 2003, 11:56:27
Job time : 87.25 secs

GenCore version 5.1.4-p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:24:32 ; Search time 58 Seconds
(without alignments)
302.472 Million cell updates/sec

Title: US-09-980-265-11

Perfect score: 20

Sequence: 1 ttacttcttagtcttagt 20

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 593429 seqs, 438583890 residues 1186858

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_MA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PC1_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PC105_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	17	85.0	659158	US-09-771-208-20	Sequence 20, Appl
2	16.8	84.0	548	US-09-924-035A-396	Sequence 396, App
3	16.8	84.0	572	US-09-864-761-12035	Sequence 12035, A
4	16.8	84.0	4920	US-08-781-986A-26	Sequence 26, Appl
5	16.4	82.0	111282	US-10-094-989-3	Sequence 3, Appl
6	15.8	79.0	397	US-09-867-701-5873	Sequence 5873, Ap
7	15.4	77.0	338	US-09-878-574-5203	Sequence 5203, Ap
8	15.4	77.0	340	US-10-025-380-568	Sequence 568, App
9	15.4	77.0	340	US-09-922-217-568	Sequence 568, App
10	15.4	77.0	340	US-09-833-263-568	Sequence 568, App
11	15.4	77.0	2000	US-09-938-842A-4252	Sequence 4252, Ap
12	15.2	76.0	207	US-09-867-701-4546	Sequence 4546, Ap
13	15.2	76.0	241	US-09-924-035A-72	Sequence 72, Appl
14	15.2	76.0	471	US-09-736-457-524	Sequence 524, App
15	15.2	76.0	471	US-09-902-941-524	Sequence 524, App
16	15.2	76.0	471	US-09-849-626-524	Sequence 524, App
17	15.2	76.0	471	US-10-017-754-524	Sequence 524, App
18	15.2	76.0	483	US-10-202-193-239	Sequence 239, App
19	15.2	76.0	711	US-09-770-149-200	Sequence 200, App

20	15.2	76.0	742	US-09-770-149-51	Sequence 51, Appl
21	15.2	76.0	742	US-09-770-149-54	Sequence 54, Appl
22	15.2	76.0	815	US-10-202-193-240	Sequence 240, App
23	15.2	76.0	995	US-09-956-004-101	Sequence 101, App
24	15.2	76.0	1244	US-09-981-353-79	Sequence 79, Appl
25	15.2	76.0	1288	US-09-925-301-596	Sequence 596, App
26	15.2	76.0	1478	US-09-938-842A-2773	Sequence 2773, Ap
27	15.2	76.0	1557	US-10-001-873-1	Sequence 1, Appl
28	15.2	76.0	2000	US-09-938-842A-4454	Sequence 4454, Ap
29	15.2	76.0	2000	US-09-938-842A-4984	Sequence 4984, Ap
30	15.2	76.0	2000	US-09-938-842A-5278	Sequence 5278, Ap
31	15.2	76.0	2000	US-09-938-842A-5298	Sequence 5298, Ap
32	15.2	76.0	2002	US-09-887-576-3	Sequence 3, Appl
33	15.2	76.0	2010	US-09-887-576-2	Sequence 2, Appl
34	15.2	76.0	2785	US-09-802-674-3	Sequence 3, Appl
35	15.2	76.0	3231	US-08-754-311B-1	Sequence 1, Appl
36	15.2	76.0	3498	US-10-091-504-2204	Sequence 2204, Ap
37	15.2	76.0	3498	US-10-091-504-2271	Sequence 2271, Ap
38	15.2	76.0	3498	US-09-764-869-2204	Sequence 2204, Ap
39	15.2	76.0	3498	US-09-764-869-2271	Sequence 2271, Ap
40	15.2	76.0	3499	US-10-091-504-2270	Sequence 2270, Ap
41	15.2	76.0	3499	US-10-091-504-2270	Sequence 2270, Ap
42	15.2	76.0	3499	US-09-764-869-2203	Sequence 2203, Ap
43	15.2	76.0	3499	US-09-764-869-2270	Sequence 2270, Ap
44	15.2	76.0	20165	US-10-199-330-7	Sequence 7, Appl
45	15.2	76.0	20165	US-10-199-334-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-09-771-208-20
Sequence 20, Application US/09771208
Patent No. US20020155564A1
GENERAL INFORMATION:
APPLICANT: MEDRANO, JUAN
APPLICANT: BRADFORD, ERIC
APPLICANT: HORVAT, SIMON
TITLE OF INVENTION: CLONING OF A HIGH-GROWTH GENE
FILE REFERENCE: 407T-923710US
CURRENT APPLICATION NUMBER: US/09/771,208
CURRENT FILING DATE: 2001-01-26
PRIORITY APPLICATION NUMBER: US 08/999,477
PRIORITY FILING DATE: 1997-12-29
NUMBER OF SEQ ID NOS: 20
SOFTWARE: PatentIn version 3.0
SEQ ID NO 20
LENGTH: 659158
TYPE: DNA
ORGANISM: Mus musculus
FEATURE:
NAME/KEY: misc.feature
LOCATION: (123459)..(123478)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc.feature
LOCATION: (602466)..(602485)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc.feature
LOCATION: (546598)..(547017)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc.feature
LOCATION: (494715)..(494814)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc.feature
LOCATION: (390865)..(391005)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc.feature
LOCATION: (346860)..(346823)
OTHER INFORMATION: n is unidentified a, c, g, or t
NAME/KEY: misc.feature
LOCATION: (317174)..(317193)
OTHER INFORMATION: n is unidentified a, c, g, or t

```
; NAME/KEY: misc_feature
; LOCATION: (280353)..(280373)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (271829)..(271848)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (183872)..(183891)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (170625)..(170645)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; LOCATION: (132680)..(132700)
; OTHER INFORMATION: n is unidentified a, c, g, or t
; NAME/KEY: misc_feature
; OTHER INFORMATION: n is a, c, g, or t
US-09-771-208-20
```

```
Query Match      85.0%; Score 17; DB 9; Length 659158;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      4 GTCTTAGTCTTAGT 20
      |||||
Db 375807 GTCTTAGTCTTAGT 375823
```

```
RESULT 2
US-09-924-035A-396/C
; Sequence 396, Application US/09924035A
; Patent No. US20020142319A1
; GENERAL INFORMATION:
; APPLICANT: GliaCh, Jrn
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 201JUS
; CURRENT APPLICATION NUMBER: US/09/924_035A
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,784
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 396
; LENGTH: 548
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(348)
; OTHER INFORMATION: n = A,T,C or G
US-09-924-035A-396
```

```
Query Match      84.0%; Score 16.8; DB 10; Length 548;
Best Local Similarity 90.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      1 TTAGTCTTAGTCTTAGT 20
      |||||
Db      172 TTAGTCTTAGTCTTAGT 153
```

```
RESULT 3
US-09-864-761-12035
; Sequence 12035, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
```

```
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 12035
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL133372.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.83
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.66
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.64
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.6
US-09-864-761-12035
```

```
Query Match      84.0%; Score 16.8; DB 10; Length 572;
Best Local Similarity 90.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      1 TTAGTCTTAGTCTTAGT 20
      |||||
Db      403 TTATTCTTAGTCTTAGT 422
```

```
RESULT 4
US-08-781-986A-26/C
; Sequence 26, Application US/08781986A
; Publication No. US20030054436A1
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 5255
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
```

STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,986A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Benson, Bob
REGISTRATION NUMBER: 30,446
REFERENCE/DOCKET NUMBER: PB248BP
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 4920 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-781-986A-26

Query Match 84.0%; Score 16.8; DB 7; Length 4920;
Best Local Similarity 90.0%; Pred. No. 1.6e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTAGT 20
|||||
DB 564 TTAATCTTACTCTTAGT 545

RESULT 5
US-10-094-989-3
; Sequence 3, Application US/10094989
; Patent No. US20020115179A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHODIESTERASE
; TITLE OF INVENTION: PROTEINS, NUCLEIC ACID MOLECULES ENCODING HUMAN
; FILE REFERENCE: CL001063DIY
; CURRENT APPLICATION NUMBER: US/10/094,989
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 09/754,250
; PRIOR FILING DATE: 2001-01-05
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 111282
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...((111282)
; OTHER INFORMATION: n = A,T,C or G
US-10-094-989-3

Query Match 82.0%; Score 16.4; DB 12; Length 111282;
Best Local Similarity 94.4%; Pred. No. 3.3e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 AGTTCTTAGTCTTAGT 20
|||||

DB 102136 AGTCTTAGTCTTAGAT 102153

RESULT 6
US-09-867-701-5873
; Sequence 5873, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5873
; LENGTH: 397
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-5873

Query Match 79.0%; Score 15.8; DB 10; Length 397;
Best Local Similarity 89.5%; Pred. No. 3.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAGTCTTAGTCTTAGT 20
|||||
DB 100 TAGTCTTAGTCTTAGAT 118

RESULT 7
US-09-878-574-5203
; Sequence 5203, Application US/09878574
; Patent No. US20020110548A1
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Thompson, Michael D.
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15401)B
; CURRENT APPLICATION NUMBER: US/09/878,574
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 09/333,535
; PRIOR FILING DATE: 1999-06-14
; NUMBER OF SEQ ID NOS: 15775
; SEQ ID NO 5203
; LENGTH: 338
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: LIB3028-029-01-B1-E5
US-09-878-574-5203

Query Match 77.0%; Score 15.4; DB 10; Length 338;
Best Local Similarity 94.1%; Pred. No. 4.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTA 17
|||||
DB 215 TTAATCTTAGTCTTA 231

RESULT 8
US-10-025-380-568
; Sequence 568, Application US/10025380
; Publication No. US20020182191A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather

```

; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole L.
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Skelky, Yasir A. W.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darick
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C14
; CURRENT APPLICATION NUMBER: US/10/025.380
; CURRENT FILING DATE: 2001-12-19
; NUMBER OF SEQ ID NOS: 1129
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 568
; LENGTH: 340
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 326
; OTHER INFORMATION: n = A,T,C or G
US-10-025-380-568
```

```

Query Match          77.0%; Score 15.4; DB 9; Length 340;
Best Local Similarity 94.1%; Pred. No. 4.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 TTAGTCTTAGTCTTA 17
Db 227 TTGGTCTTAGTCTTA 243
```

```

RESULT 9
US-09-922-217-568
; Sequence 568, Application US/09922217
; Patent No. US20020076414A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Jiang, Yugu
; APPLICANT: Smith, Carole Lynn
; APPLICANT: King, Gordon E.
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
; FILE REFERENCE: 210121.471C13
; CURRENT APPLICATION NUMBER: US/09/922.217
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 568
; LENGTH: 340
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 326
; OTHER INFORMATION: n = A,T,C or G
US-09-922-217-568
```

```

Query Match          77.0%; Score 15.4; DB 10; Length 340;
Best Local Similarity 94.1%; Pred. No. 4.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 TTAGTCTTAGTCTTA 17
Db 227 TTGGTCTTAGTCTTA 243
```

```

RESULT 10
US-09-833-263-568
; Sequence 568, Application US/09833263
; Patent No. US2002010547A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Aljun
; APPLICANT: Clapper, Jonathan D.
; APPLICANT: Stolk, John A.
; APPLICANT: Meagher, Madeleine J.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; FILE REFERENCE: 210121.471C12
; CURRENT APPLICATION NUMBER: US/09/833.263
; CURRENT FILING DATE: 2001-04-10
; NUMBER OF SEQ ID NOS: 1093
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 568
; LENGTH: 340
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(340)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-263-568
```

```

Query Match          77.0%; Score 15.4; DB 10; Length 340;
Best Local Similarity 94.1%; Pred. No. 4.6e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 1 TTAGTCTTAGTCTTA 17
Db 227 TTGGTCTTAGTCTTA 243
```

```

RESULT 11
US-09-938-842A-4252
; Sequence 4252, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCDP1300-3
; CURRENT APPLICATION NUMBER: US/09/938.842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 4252
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-4252
```

```

Query Match          77.0%; Score 15.4; DB 9; Length 2000;
Best Local Similarity 94.1%; Pred. No. 5.7e+02;
```


Matches 16: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTCTTA 17
|||||
Db 1367 TTAGTCTTAGTCTTCTTA 1383

RESULT 12
US-09-867-701-4546/c
; Sequence 4546, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4546
; LENGTH: 207
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-867-701-4546

Query Match 76.0%; Score 15.2; DB 10; Length 207;
Best Local Similarity 85.0%; Pred. No. 5.3e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTCTTCTT 20
|||||
Db 113 TTGCTTCTTAGTCTTCTTCTT 94

RESULT 13
US-09-924-035A-72/c
; Sequence 72, Application US/09924035A
; Patent No. US20020142319A1
; GENERAL INFORMATION:
; APPLICANT: Grilach, Jrm
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2011US
; CURRENT APPLICATION NUMBER: US/09/924,035A
; CURRENT FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: US 60/148,784
; PRIOR FILING DATE: 1999-08-13
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 72
; LENGTH: 241
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-924-035A-72

Query Match 76.0%; Score 15.2; DB 10; Length 241;
Best Local Similarity 85.0%; Pred. No. 5.4e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTCTTCTT 20
|||||
Db 87 TTACTTCTTAGTCTTCTTCTT 68

RESULT 14
US-09-736-457-524/c
; Sequence 524, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong

; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darlick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 524
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-524

Query Match 76.0%; Score 15.2; DB 9; Length 471;
Best Local Similarity 85.0%; Pred. No. 5.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTCTTCTT 20
|||||
Db 190 TTAGCTTCTTAGTCTTCTTCTT 171

RESULT 15
US-09-902-941-524/c
; Sequence 524, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tonglong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darlick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 524
; LENGTH: 471
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-524

Query Match 76.0%; Score 15.2; DB 9; Length 471;
Best Local Similarity 85.0%; Pred. No. 5.9e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTCTTAGTCTTCTTCTT 20
|||||
Db 190 TTAGCTTCTTAGTCTTCTTCTT 171

Search completed: April 6, 2003, 12:12:31
Job time : 212 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 10:29:22 ; Search time 168.5 seconds
(without alignments)
1912.600 Million cell updates/sec

Title: US-09-980-265-12
Perfect score: 20
Sequence: 1 ttagttgttagttattgctt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 60155 seqs, 8056825618 residues

Total number of hits satisfying chosen parameters: 120310

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl_HTG:*
1: gb_hgtc:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
C 1	19	95.0	195444	1	AC107399 Homo sapi
C 2	18.4	92.0	5421	1	AC019958 Drosophi
C 3	18.4	92.0	38069	1	AC017255 Drosophi
C 4	18.4	92.0	44765	1	AC116467 Mus muscu
C 5	18.4	92.0	80679	1	AC103576 Rattus no
C 6	18.4	92.0	95180	1	AC109057 Rattus no
C 7	18.4	92.0	106651	1	AC106466 Rattus no
C 8	18.4	92.0	150326	1	AC128956 Rattus no
C 9	18.4	92.0	153324	1	AC129017 Mus muscu
C 10	18.4	92.0	164552	1	AC111380 Rattus no
C 11	18.4	92.0	176400	1	AC113369 Homo sapi
C 12	18.4	92.0	178712	1	AC115176 Rattus no
C 13	18.4	92.0	178970	1	AC012589 Homo sapi
C 14	18.4	92.0	201132	1	AC122335 Mus muscu
C 15	18.4	92.0	241585	1	AL844536 Mus muscu
C 16	17.4	87.0	20733	1	AC019829 Drosophi
C 17	17.4	87.0	60625	1	AC020090 Drosophi
C 18	17.4	87.0	64061	1	AC100401 Mus muscu
C 19	17.4	87.0	92053	1	AC017270 Drosophi
C 20	17.4	87.0	154354	1	AC017270 Drosophi
C 21	17.4	87.0	167683	1	AC079251 Homo sapi
C 22	17.4	87.0	171272	1	AC084707 Homo sapi
C 23	17.4	87.0	171457	1	AC112568 Rattus no
C 24	17.4	87.0	174877	1	AP004818 Rattus no
C 25	17.4	87.0	177741	1	AC118176 Rattus no
C 26	17.4	87.0	180485	1	AC099106 Rattus no
C 27	17.4	87.0	183450	1	AC119322 Rattus no
C 28	17.4	87.0	186385	1	AC026821 Homo sapi
C 29	17.4	87.0	186542	1	AL844197 Danto rer
C 30	17.4	87.0	193039	1	AC090690 Homo sapi
C 31	17.4	87.0	195600	1	AC079094 Homo sapi
C 32	17.4	87.0	209572	1	AC099615 Mus muscu

C 33	17.4	87.0	209790	1	AL512373 Homo sapi
C 34	17.4	87.0	216464	1	AC102540 Mus muscu
C 35	17.4	87.0	247102	1	AC123854 Mus muscu
C 36	17.4	87.0	267052	1	AC122872 Mus muscu
C 37	17.4	85.0	23154	1	AC115602 Dictyoste
C 38	17	85.0	56759	1	AC025379 Homo sapi
C 39	17	85.0	68559	1	AC031981 Homo sapi
C 40	17	85.0	160294	1	AC025621 Homo sapi
C 41	17	85.0	166349	1	AC013707 Homo sapi
C 42	17	85.0	168208	1	AC024359 Homo sapi
C 43	17	85.0	214271	1	AC090737 Homo sapi
C 44	16.8	84.0	8659	1	AC117270 Dictyoste
C 45	16.8	84.0	14921	1	AC015356 Drosophi

ALIGNMENTS

RESULT 1
AC107399/c
LOCUS
DEFINITION Homo sapiens chromosome 4 clone RP11-755N14, WORKING DRAFT
ACCESSION AC107399
VERSION AC107399.5 GI:21450604
KEYWORDS
SOURCE HTG: HTGS_PHASE1, HTGS_DRAFT, HTGS_ACTIVEPIN.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
Waterston, R.H.
1 (bases 1 to 195444)
2 (bases 1 to 195444)
Unpublished
The sequence of Homo sapiens clone
Waterston, R.H.
Direct Submission
Submitted (19-JUN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
3 (bases 1 to 195444)
Waterston, R.H.
Direct Submission
Submitted (20-JUN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jun 18, 2002 this sequence version replaced gi:21426762.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
Project Information
Center project name: H_NH0755N14

Summary Statistics

Sequencing vector: M13, 0%
Sequencing vector: Plasmid, 100%
Chemistry: Dye-terminator Big Dye, 100% of reads
Chemistry: Dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 194367 bases at least Q40
Consensus quality: 194908 bases at least Q30
Consensus quality: 195506 bases at least Q20
Insert size: 211000; agarose-fp
Insert size: 196567; sum-of-ctrls
Quality coverage: 7.60 in Q20 bases; sum-of-ctrls
Quality coverage: 6.70 in Q20 bases; sum-of-ctrls

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1243: contig of 1243 bp in length
* 1244 1343: gap of unknown length
* 1344 94590: contig of 93247 bp in length
* 94591 94690: gap of unknown length
* 94691 195444: contig of 100754 bp in length.
Location/Qualifiers
1. 195444

FEATURES
source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="4"
/clone="RP11-755N14"
1. 1243
/note="assembly_name:Contig39"
1344. 94590
/note="assembly_name:Contig42"
94691. 195444
/note="assembly_name:Contig43
clone_end:T7
vector_side:right"

BASE COUNT 64030 a 32571 c 33465 g 65177 t 201 others
ORIGIN

Query Match 95.0%; Score 19; DB 1; Length 195444;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 TTACTGTTAGTTATTTAGT 19
Db 28807 TTAGTTGTTAGTTATTTAGT 28789

RESULT 2
AC019958/c

LOCUS AC019958 5421 bp DNA linear HTG 03-JAN-2000
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
AC019958
AC019958.1 GI:6664939
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 5421)
Adams,M. and Venter,J.C.
Direct Submission
Submitted (30-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

This sequence was identified as CDW:10211530 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

COMMENT

Location/Qualifiers
1. 5421
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

FEATURES
source

BASE COUNT 1558 a 1150 c 1146 g 1567 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 5421;
Best Local Similarity 95.0%; Pred. No. 14;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTACTGTTAGTTATTTAGT 20
TTTTTTTTTTTTTTTTTTTT

Db 428 TTACTGTTAGTTCTTAGTT 409

RESULT 3
AC017255

LOCUS AC017255 38069 bp DNA linear HTG 09-DEC-1999
DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***, in ordered
pieces.
AC017255
AC017255.1 GI:6553731
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 38069)
Adams,M. and Venter,J.C.
Direct Submission
Submitted (09-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA

REFERENCE
AUTHORS
TITLE
JOURNAL

This sequence was identified as CDW:10210065 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

Location/Qualifiers
1. 38069
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"

FEATURES
source

BASE COUNT 12432 a 7322 c 7198 g 11117 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 38069;
Best Local Similarity 95.0%; Pred. No. 17;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTACTGTTAGTTATTTAGT 20
Db 2372 TTAGTTATTTAGTTATTTAGT 2391

RESULT 4
AC116467

LOCUS AC116467 44765 bp DNA linear HTG 28-MAR-2002
DEFINITION Mus musculus clone RP23-165115, LOW-PASS SEQUENCE SAMPLING.
AC116467
AC116467.1 GI:19774573
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Mus musculus.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 44765)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Mus musculus, clone RP23-165115
Unpublished
2 (bases 1 to 44765)

REFERENCE
AUTHORS
TITLE
JOURNAL

Anderson,S., Barna,N., Bastien,V., Bloom,T., Bouuslavsky,L.,
Boukhalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J.,
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,
Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Johnson,R., Jones,C.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Lahocque,K., Lamazares,R.,
Landers,T., Lehoczek,J., Levine,R., Lindblad-Toh,K., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Menus,L.,
Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,

TITLE
JOURNAL
COMMENT

Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Sudramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (28-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/rw/RepeatMasker.html>

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L23458

Center clone name: 165_I_15

* NOTE: This record contains 58 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
* 655 754: gap of 100 bp in length
* 755 1457: contig of 703 bp in length
* 1458 1557: gap of 100 bp
* 1558 2210: contig of 653 bp in length
* 2211 2310: gap of 100 bp
* 2311 2988: contig of 678 bp in length
* 2989 3088: gap of 100 bp
* 3089 3776: contig of 688 bp in length
* 3777 3876: gap of 100 bp
* 3877 4535: contig of 659 bp in length
* 4536 4635: gap of 100 bp
* 4636 5329: contig of 694 bp in length
* 5330 5429: gap of 100 bp
* 5430 6104: contig of 675 bp in length
* 6105 6204: gap of 100 bp
* 6205 6879: contig of 675 bp in length
* 6880 6979: gap of 100 bp
* 6980 7669: contig of 690 bp in length
* 7670 7769: gap of 100 bp
* 7770 8440: contig of 671 bp in length
* 8441 8540: gap of 100 bp
* 8541 9219: contig of 679 bp in length
* 9220 9319: gap of 100 bp
* 9320 9990: contig of 671 bp in length
* 9991 10090: gap of 100 bp
* 10091 10746: contig of 656 bp in length
* 10747 10846: gap of 100 bp
* 10847 11497: contig of 651 bp in length
* 11498 11597: gap of 100 bp
* 11598 12303: contig of 706 bp in length
* 12304 12403: gap of 100 bp
* 12404 13046: contig of 643 bp in length
* 13047 13146: gap of 100 bp
* 13147 13834: contig of 688 bp in length
* 13835 13934: gap of 100 bp
* 13935 14610: contig of 676 bp in length
* 14611 14710: gap of 100 bp
* 14711 15396: contig of 686 bp in length

15397 15496: gap of 100 bp
15497 16152: contig of 656 bp in length
16153 16252: gap of 100 bp
16253 16949: contig of 697 bp in length
16950 17049: gap of 100 bp
17050 17720: contig of 671 bp in length
17721 17820: gap of 100 bp
17821 18524: contig of 704 bp in length
18525 18624: gap of 100 bp
18625 19335: contig of 711 bp in length
19336 19435: gap of 100 bp
19436 20126: contig of 691 bp in length
20127 20226: gap of 100 bp
20227 20906: contig of 680 bp in length
20907 21006: gap of 100 bp
21007 21687: contig of 681 bp in length
21688 21787: gap of 100 bp
21788 22456: contig of 669 bp in length
22457 22556: gap of 100 bp
22557 23229: contig of 673 bp in length
23230 23329: gap of 100 bp
23330 23329: gap of 100 bp
24012 24111: gap of 100 bp
24112 24758: contig of 647 bp in length
24759 24858: gap of 100 bp
24859 25531: contig of 673 bp in length
25532 25631: gap of 100 bp
25632 26304: contig of 673 bp in length
26305 26404: gap of 100 bp
26405 27034: contig of 630 bp in length
27035 27134: gap of 100 bp
27135 27841: contig of 707 bp in length
27842 27941: gap of 100 bp
27942 28584: contig of 643 bp in length
28585 28684: gap of 100 bp
28685 29346: contig of 662 bp in length
29347 29446: gap of 100 bp
29447 30136: contig of 690 bp in length
30137 30236: gap of 100 bp
30237 30906: contig of 670 bp in length
30907 31006: gap of 100 bp
31007 31686: contig of 680 bp in length
31687 31786: gap of 100 bp
31787 32456: contig of 670 bp in length
32457 32556: gap of 100 bp
32557 33196: contig of 640 bp in length
33197 33286: gap of 100 bp
33297 33977: contig of 681 bp in length
33978 34077: gap of 100 bp
34078 34704: contig of 627 bp in length
34705 34804: gap of 100 bp
34805 35481: contig of 677 bp in length
35482 35581: gap of 100 bp
35582 36251: contig of 670 bp in length
36252 36351: gap of 100 bp
36352 37049: contig of 698 bp in length
37050 37149: gap of 100 bp
37150 37797: contig of 648 bp in length
37798 37897: gap of 100 bp
37898 38601: contig of 704 bp in length
38602 38701: gap of 100 bp
38702 39324: contig of 623 bp in length
39325 39424: gap of 100 bp
39425 40112: contig of 688 bp in length
40113 40212: gap of 100 bp
40213 40892: contig of 680 bp in length
40893 40992: gap of 100 bp
40993 41656: contig of 664 bp in length
41657 41756: gap of 100 bp
41757 42447: contig of 691 bp in length
42448 42547: gap of 100 bp
42548 43206: contig of 659 bp in length
43207 43306: gap of 100 bp

FEATURES * 43307 43984: contig of 678 bp in length
* 43985 44084: gap of 100 bp
* 44085 44765: contig of 681 bp in length.
location/Qualifiers
1. 44765
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_id="RPC1-23 Female Mouse BAC"
BASE COUNT 11666 a 7677 c 7735 g 11806 t 5881 others
ORIGIN

Query Match 92.0% Score 18.4; DB 1; Length 44765;
Best Local Similarity 95.0%; Pred. No. 17;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTACTGTGTTACTTACTT 20
||||||| |||||||
Db 4039 TTACTGTGTTACTTACTT 4058

RESULT 5
AC103576
LOCUS
DEFINITION Rattus norvegicus clone CH230-63M13, *** SEQUENCING IN PROGRESS
AC103576 80679 bp DNA linear HTG 12-JUL-2002
AC103576
VERSION AC103576.4 GI:21729709
KEYWORDS HTG; HTGS PHASEL.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 80679)
Mizny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Aishrooks,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbada,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaune,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Hollway,C., Hollins,B.,
Homsli,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacksom,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsom,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,S., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokewo,S., Ogih,M., Okunou,G.,
Otagunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Severy,G.,
Sheehy,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,Y., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wlezyk,R., Wooden,C., Worley,K.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 11, 2002 this sequence version replaced gi:18846139.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GIGY
Center clone name: CH230-63M13
----- Summary Statistics
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 33297 bases at least Q40
Consensus quality: 35946 bases at least Q30
Consensus quality: 38652 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 47 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1171: contig of 1171 bp in length
1172 1271: gap of unknown length
1272 2868: contig of 1597 bp in length
2869 2968: gap of unknown length
2969 4131: contig of 1163 bp in length
4132 4232: gap of unknown length
4233 5312: contig of 1081 bp in length
5313 5412: gap of unknown length
5413 6870: contig of 1458 bp in length
6871 6970: gap of unknown length
6971 7992: contig of 1021 bp in length
7993 8091: gap of unknown length
8092 9125: contig of 1035 bp in length
9126 9226: gap of unknown length
9227 10468: contig of 1242 bp in length
10469 10568: gap of unknown length
10569 11619: contig of 1051 bp in length
11620 11719: gap of unknown length
11720 12894: contig of 1175 bp in length
12895 12994: gap of unknown length
12995 14178: contig of 1184 bp in length
14179 14278: gap of unknown length
14279 15615: contig of 1337 bp in length
15616 15715: gap of unknown length
15716 17016: contig of 1301 bp in length
17017 17116: gap of unknown length
17118 18233: contig of 1117 bp in length
18234 18333: gap of unknown length
18334 19665: contig of 1332 bp in length
19666 21151: gap of unknown length
21152: contig of 1386 bp in length

* 21152 21251: gap of unknown length
* 21252 22739: contig of 1488 bp in length
* 22740 22839: gap of unknown length
* 22840 23909: contig of 1070 bp in length
* 23910 24009: gap of unknown length
* 24010 25502: contig of 1493 bp in length
* 25503 25602: gap of unknown length
* 25603 26796: contig of 1194 bp in length
* 26797 26896: gap of unknown length
* 26897 28120: contig of 1224 bp in length
* 28121 28220: gap of unknown length
* 28221 29944: contig of 1724 bp in length
* 29945 30044: gap of unknown length
* 30045 32238: contig of 2194 bp in length
* 32239 32338: gap of unknown length
* 32339 33569: contig of 1231 bp in length
* 33570 33669: gap of unknown length
* 33670 33660: contig of 1891 bp in length
* 33661 35660: gap of unknown length
* 35661 36788: contig of 1128 bp in length
* 36789 36888: gap of unknown length
* 36889 38239: contig of 1351 bp in length
* 38240 38339: gap of unknown length
* 38340 39574: contig of 1235 bp in length
* 39575 39674: gap of unknown length
* 39675 40719: contig of 1045 bp in length
* 40720 40819: gap of unknown length
* 40820 42444: contig of 1625 bp in length
* 42445 42544: gap of unknown length
* 42545 44516: contig of 1972 bp in length
* 44517 44616: gap of unknown length
* 44617 46392: contig of 1776 bp in length
* 46393 46492: gap of unknown length
* 46493 47848: contig of 1356 bp in length
* 47849 47948: gap of unknown length
* 47949 49380: contig of 1432 bp in length
* 49381 49480: gap of unknown length
* 49481 50992: contig of 1512 bp in length
* 50993 51092: gap of unknown length
* 51093 53314: contig of 2122 bp in length
* 53315 53314: gap of unknown length
* 53315 54686: contig of 1372 bp in length
* 54687 54786: gap of unknown length
* 54787 56476: contig of 1690 bp in length
* 56477 56576: gap of unknown length
* 56577 58624: contig of 2048 bp in length
* 58625 58724: gap of unknown length
* 58725 61311: contig of 2587 bp in length
* 61312 61411: gap of unknown length
* 61412 63407: contig of 1996 bp in length
* 63408 63507: gap of unknown length
* 63508 65412: contig of 1905 bp in length
* 65413 65512: gap of unknown length
* 65513 68121: contig of 2609 bp in length
* 68122 68221: gap of unknown length
* 68222 71194: contig of 2973 bp in length
* 71195 71294: gap of unknown length
* 71295 73654: contig of 2360 bp in length
* 73655 73754: gap of unknown length
* 73755 76459: contig of 2705 bp in length
* 76460 76559: gap of unknown length
* 76560 80679: contig of 4120 bp in length.

FEATURES

Location/Qualifiers

1. 80679
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-63N13"
BASE COUNT 17902 a 18300 c 17407 g 18485 t 8585 others
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 80679;
Best Local Similarity 95.0%; Pred. No. 18;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTTATAGTT 20
||||| |||||||||
Db 41343 TTAGTATTAGTTATAGTT 41362

RESULT 6
AC109057/c
LOCUS
DEFINITION
AC109057 95180 bp DNA linear HTG 13-JUL-2002
Rattus norvegicus clone CH230-126F19, *** SEQUENCING IN PROGRESS
***48 unordered pieces.
AC109057
AC109057.3 GI:21737841
VERSION
KEYWORDS
HTG; HTGS; PHASE1.
SOURCE
ORGANISM
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Ratus.

REFERENCE

1 (bases 1 to 95180)
Muzny,D.M., Adams,C., Adio-Oduola,B., Alt-osman,F.R., Allen,C.,
Alshrocks,S.L., Amaratunga,H.C., Are,J.R., Ayala,M., Banks,T.,
Barbaita,U., Benton,J., Binsage,K., Blankenburg,K., Bonnin,D.,
Bouck,U., Bowle,S., Brlewa,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroli,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Correll,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
Homsli,F., Howard,S., Huber,J., Huijk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kratovic,U., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Lousseng,H.,
Lozada,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogun,M., Okwono,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,D., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoshari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umsani,K., Vasquez,L., Vera,V., Villalón,D., Vltson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Welnstock,G. and Gibbs,R.

TITLE

Direct Submission

JOURNAL

Unpublished

AUTHORS

Morley,K.C.

JOURNAL

Submitted (03-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

REFERENCE

3 (bases 1 to 95180)
Morley,K.C.

AUTHORS

Direct Submission

JOURNAL

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One

COMMENT

Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced g1:18846778.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GPV6
Center clone name: CH230-126F19
----- Summary Statistics
Sequencing vector: plasmid
Chemistry: Dye-terminator Big Dye 1000 of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 59597 bases at least Q40
Consensus quality: 63979 bases at least Q30
Consensus quality: 67799 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 48 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1114: contig of 1114 bp in length
* 1115 1214: gap of unknown length
* 1215 2810: contig of 1596 bp in length
* 2811 2910: gap of unknown length
* 2911 3928: contig of 1018 bp in length
* 3929 4028: gap of unknown length
* 4029 5148: contig of 1120 bp in length
* 5149 5248: gap of unknown length
* 5249 6848: contig of 1600 bp in length
* 6849 6949: gap of unknown length
* 6949 8234: contig of 1306 bp in length
* 8235 8354: gap of unknown length
* 8355 9642: contig of 1288 bp in length
* 9643 9742: gap of unknown length
* 9743 11399: contig of 1657 bp in length
* 11400 11499: gap of unknown length
* 11500 13106: contig of 1607 bp in length
* 13107 13206: gap of unknown length
* 13207 15257: contig of 2051 bp in length
* 15258 15357: gap of unknown length
* 15358 16980: contig of 1623 bp in length
* 16981 17080: gap of unknown length
* 17081 18685: contig of 1605 bp in length
* 18686 18785: gap of unknown length
* 18786 20777: contig of 1892 bp in length
* 20778 22009: gap of unknown length
* 22010 22109: contig of 1232 bp in length
* 22110 23662: gap of unknown length
* 23663 23762: gap of unknown length
* 23763 25624: contig of 1862 bp in length
* 25625 25724: gap of unknown length
* 25725 26730: contig of 1006 bp in length
* 26731 26830: gap of unknown length
* 26831 28357: contig of 1527 bp in length
* 28358 28457: gap of unknown length
* 28458 30373: contig of 1916 bp in length
* 30374 31986: contig of 1513 bp in length
* 31987 32086: gap of unknown length
* 32087 33490: contig of 1404 bp in length
* 33491 33590: gap of unknown length
* 33591 34980: contig of 1390 bp in length
* 34981 35080: gap of unknown length
* 35081 36720: contig of 1640 bp in length

36721 36820: gap of unknown length
* 36821 38004: contig of 1184 bp in length
* 38005 38104: gap of unknown length
* 38105 39550: contig of 1446 bp in length
* 39551 39651: gap of unknown length
* 39652 40752: contig of 1001 bp in length
* 40753 43081: gap of unknown length
* 43082 43181: gap of unknown length
* 43182 44469: contig of 1288 bp in length
* 44470 44569: gap of unknown length
* 44570 45876: contig of 1307 bp in length
* 45877 45976: gap of unknown length
* 45977 48240: contig of 2264 bp in length
* 48241 48340: gap of unknown length
* 48341 50121: contig of 1781 bp in length
* 50122 50221: gap of unknown length
* 50222 51853: contig of 1632 bp in length
* 51854 51953: gap of unknown length
* 51954 53849: contig of 1896 bp in length
* 53850 53949: gap of unknown length
* 53950 56898: contig of 2949 bp in length
* 56899 56998: gap of unknown length
* 56999 58622: contig of 1624 bp in length
* 58623 58722: gap of unknown length
* 58723 60532: contig of 1810 bp in length
* 60533 60632: gap of unknown length
* 60633 61891: contig of 1259 bp in length
* 61892 61991: gap of unknown length
* 61992 64149: contig of 2158 bp in length
* 64150 64250: gap of unknown length
* 64250 66117: contig of 1868 bp in length
* 66118 66217: gap of unknown length
* 66218 67638: contig of 1421 bp in length
* 67639 67738: gap of unknown length
* 67739 70700: contig of 2962 bp in length
* 70701 70800: gap of unknown length
* 70801 73188: contig of 2388 bp in length
* 73189 73288: gap of unknown length
* 73289 75650: contig of 2371 bp in length
* 75650 75759: gap of unknown length
* 75760 78787: contig of 3028 bp in length
* 78788 78887: gap of unknown length
* 78888 82264: contig of 3377 bp in length
* 82265 82365: gap of unknown length
* 82366 84919: contig of 2555 bp in length
* 84920 85019: gap of unknown length
* 85020 90450: contig of 5431 bp in length
* 90451 90550: gap of unknown length
* 90551 95180: contig of 4630 bp in length.
Location/Qualifiers
1. 95180
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-126F19"
BASE COUNT 27139 a 18257 c 18189 g 25171 t 6424 others
ORIGIN
Query Match 92.0%; Score 18.4; DB 1; Length 95180;
Best Local Similarity 95.0%; Pred. No. 18;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTACTGTGTTACTGTTACTGTT 20
Db 53187 TTCGTTGTTACTGTTACTGTT 53168
RESULT 7
AC106466/c AC106466 106651 bp DNA linear HTG 13-JUL-2002
LOCUS Rattus norvegicus clone CH230-211B18, *** SEQUENCING IN PROGRESS
DEFINITION *** 45 unordered pieces.
ACCESSION AC106466

53750 53849: gap of unknown length
53850 55748: contig of 1899 bp in length
55749 55848: gap of unknown length
55849 59788: contig of 3940 bp in length
59789 59888: gap of unknown length
59889 62643: contig of 2761 bp in length
62650 62749: gap of unknown length
62750 65930: contig of 3181 bp in length
65931 66030: gap of unknown length
66031 68484: contig of 2454 bp in length
68485 68584: gap of unknown length
68585 70599: contig of 1925 bp in length
70510 70609: gap of unknown length
70610 73606: contig of 2997 bp in length
73607 76448: contig of 2742 bp in length
76449 76548: gap of unknown length
76549 79392: contig of 2844 bp in length
79393 79492: gap of unknown length
79493 82319: contig of 2827 bp in length
82320 82419: gap of unknown length
82420 85792: contig of 3373 bp in length
85793 85892: gap of unknown length
85893 88433: contig of 2541 bp in length
88434 88533: gap of unknown length
88534 92245: contig of 3712 bp in length
92246 92345: gap of unknown length
92346 96997: contig of 4652 bp in length
96998 97097: gap of unknown length
97098 102010: contig of 4913 bp in length
102011 102110: gap of unknown length
102111 106651: contig of 4541 bp in length.
Location/Qualifiers
1. 106651
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-211B18"

BASE COUNT 29303 a 19724 c 19632 g 30261 t 7731 others
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 106651;
Best Local Similarity 95.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTACTGTCTAGTATTAGCTT 20
||||| |||||||||
Db 83573 TTACTGTCTAGTATTAGCTT 83554

RESULT 8
AC128956
LOCUS
DEFINITION
Rattus norvegicus clone CH230-500P2, *** SEQUENCING IN PROGRESS
***, 38 unordered pieces.
AC128956 150326 bp DNA linear HTG 24-JUL-2002
AC128956.1 GI:21953705
HTG: HTGS_PHASE1.
KEYWORDS
SOURCE
Rattus norvegicus.
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 150326)

REFERENCE
AUTHORS
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbieri,J., Benton,J., Bimge,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhai,C., Escoto,M.,
Fall,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabali,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guenara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A., Hernandez,J.,
Hernandez,O., Hodgeson,A., Hognes,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Huylk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., Kling,L., Korvah,J., Kovar,C.,
Kratovic,J., Kurehli,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,M., Louised,H.,
Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mel,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabath,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okunnu,G.,
Oragunye,N., Oyedele,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabot,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansley,J., Taylor,C., Taylor,T., Teitford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wlezyk,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 150326)
Direct Submission
Submitted (24-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: KBPH
Center clone name: CH230-500P2

Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 110267 bases at least Q40
Consensus quality: 116757 bases at least Q30
Consensus quality: 121521 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 38 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 1047: contig of 1047 bp in length
* 1048 1147: gap of unknown length
* 1148 2218: contig of 1071 bp in length
* 2219 3360: gap of unknown length
* 3361 3460: contig of 1042 bp in length
* 3461 5074: gap of unknown length
* 5075 5174: contig of 1614 bp in length
* 5175 6294: gap of unknown length
* 6294: contig of 1120 bp in length

```
* 6295 6394: gap of unknown length
* 6395 8169: contig of 1775 bp in length
* 8170 8269: gap of unknown length
* 8270 9298: contig of 1029 bp in length
* 9299 9398: gap of unknown length
* 9399 10923: contig of 1525 bp in length
* 10924 11023: gap of unknown length
* 11024 12507: contig of 1484 bp in length
* 12508 12607: gap of unknown length
* 12608 14841: contig of 2234 bp in length
* 14842 14941: gap of unknown length
* 14942 17290: contig of 2349 bp in length
* 17291 17390: gap of unknown length
* 17391 18790: contig of 1400 bp in length
* 18791 18890: gap of unknown length
* 18891 20400: contig of 1510 bp in length
* 20401 20500: gap of unknown length
* 20501 22551: contig of 2051 bp in length
* 22552 22651: gap of unknown length
* 22652 24396: contig of 1745 bp in length
* 24397 24496: gap of unknown length
* 24497 26180: contig of 1684 bp in length
* 26181 26280: gap of unknown length
* 26281 28087: contig of 1807 bp in length
* 28088 28187: gap of unknown length
* 28188 30547: contig of 2360 bp in length
* 30548 30647: gap of unknown length
* 30648 33337: contig of 2690 bp in length
* 33338 33437: gap of unknown length
* 33438 36007: contig of 2570 bp in length
* 36008 36107: gap of unknown length
* 36108 39900: contig of 3793 bp in length
* 39901 40000: gap of unknown length
* 40001 43740: contig of 3740 bp in length
* 43741 43840: gap of unknown length
* 43841 47274: contig of 3434 bp in length
* 47275 47374: gap of unknown length
* 47375 52052: contig of 4678 bp in length
* 52053 52152: gap of unknown length
* 52153 55896: contig of 3744 bp in length
* 55897 55996: gap of unknown length
* 55997 61630: contig of 5634 bp in length
* 61631 61730: gap of unknown length
* 61731 65336: contig of 3606 bp in length
* 65337 65436: gap of unknown length
* 65437 69824: contig of 4388 bp in length
* 69825 69924: gap of unknown length
* 69925 74153: contig of 4229 bp in length
* 74154 74253: gap of unknown length
* 74254 80624: contig of 6371 bp in length
* 80625 85103: gap of unknown length
* 85104 85203: contig of 4379 bp in length
* 85204 90043: gap of unknown length
* 90044 90143: gap of unknown length
* 90144 94288: contig of 4145 bp in length
* 94289 94388: gap of unknown length
* 94389 99899: contig of 5511 bp in length
* 99900 99999: gap of unknown length
* 100000 107750: contig of 7751 bp in length
* 107751 107850: gap of unknown length
* 107851 114037: contig of 6187 bp in length
* 114038 114137: gap of unknown length
* 114138 132540: contig of 18403 bp in length
* 132541 132641: gap of unknown length
* 132641 150326: contig of 17686 bp in length.
```

FEATURES

Location/Qualifiers

1..150326

/organism="Rattus norvegicus"

/db_xref="taxon:10116"

/clone="CH230-500b2"

BASE COUNT 36075 a 34635 c 33943 g 36814 t 8859 others
ORIGIN

Query Match 92.0%; Score 18.4; DB 1; Length 150326;
Best Local Similarity 95.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGTT 20

Db 114899 TTAGTTGTTAGTTAGTTAGTT 114918

RESULT 9

AC129017/c

LOCUS

DEFINITION

AC129017

AC129017.1 GI:21954045

KEYWORDS

HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE

Mus musculus

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Contact: submissions@wustl.edu
Project Information
Center project name: M.BB041A24

----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 148742 bases at least Q40
Consensus quality: 149676 bases at least Q30
Consensus quality: 150242 bases at least Q20
Insert size: 152557; sum-of-ctgts
Quality coverage: 6.72 in Q20 bases; agarose-fp
Quality coverage: 6.60 in Q20 bases; sum-of-ctgts

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1133: contig of 1133 bp in length
* 1233: gap of unknown length
* 1234 4269: contig of 3036 bp in length
* 4270 4369: gap of unknown length
* 4370 8447: contig of 4078 bp in length
* 8448 8547: gap of unknown length
* 8548 20096: contig of 11449 bp in length
* 20097 20196: gap of unknown length
* 20197 40259: contig of 20063 bp in length
* 40260 40359: gap of unknown length
* 40360 63542: contig of 23183 bp in length

FEATURES

Source

1.152324
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="X"X
/clone="RP24-41A24"

misc_feature
1.1133
/note="assembly_name:Contig13"

misc_feature
1234.4269
/note="assembly_name:Contig14"

misc_feature
4370.8447
/note="assembly_name:Contig15"

misc_feature
vector_side:left"

misc_feature
8348.20096
/note="assembly_name:Contig16"

misc_feature
20197.40259
/note="assembly_name:Contig17"

misc_feature
40360.63542
/note="assembly_name:Contig18"

misc_feature
63643.85165
/note="assembly_name:Contig19"

misc_feature
clone_end:T7
vector_side:left"

misc_feature
85266.118186
/note="assembly_name:Contig20"

misc_feature
118287.152324
/note="assembly_name:Contig21"

BASE COUNT 49091 a 27882 c 28512 g 45991 t 848 others

ORIGIN

Query Match 92.0% Score 18.4: DB 1: Length 152324:
Best Local Similarity 95.0%: Pred. No. 19:
Matches 19: Conservative 0: Mismatches 1: Indels 0: Gaps 0:

Qy 1 TTAGTGTAGTATTAGT 20
|||||TTTTTTTTTTTT

Db 28012 TTAGTGTAGTATTAGT 27993

RESULT 10
AC11380
LOCUS
DEFINITION Rattus norvegicus clone CH230-130H20, *** SEQUENCING IN PROGRESS
AC11380
AC11380.2 GI:21735908
VERSION
KEYWORDS HTG: HTGS PHASE1.
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 164552)

REFERENCE
AUTHORS Muzny,D.M., Adams,C., Adio-oduola,B., All-osman,F.R., Allen,C.,
Alsdbrook,S.L., Amaralunge,H.C., Aye,J.R., Ayele,M., Banks,T.,
Barbieri,J., Benton,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowler,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carton,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Douchlatte,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,

Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolyet,S., Joudah,S.,
Karlisson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kralovic,J., Kureshi,A., Landry,N., Deal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,M., Louised,H.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Merindale,A., Martinez,E.,
Massey,E., Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabadi,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,S., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokwenkwo,S., Oguh,M., Okwuonu,G.,
Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pui,L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Saverly,G.,
Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I.,
Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,Y., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.

Direct Submission
Unpublished
2 (bases 1 to 164552)
Worley,K.C.

REFERENCE
AUTHORS Morley,K.C.

JOURNAL
TITLE Direct Submission

REFERENCE
AUTHORS Morley,K.C.

JOURNAL
TITLE Direct Submission

REFERENCE
AUTHORS Morley,K.C.

JOURNAL
TITLE Direct Submission

COMMENT

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

Project Information
Center project name: GMDJ
Center clone name: CH230-130H20

Summary Statistics
Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 106093 bases at least Q40
Consensus quality: 110397 bases at least Q30
Consensus quality: 114114 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 58 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1002: contig of 1002 bp in length
* 1003 1102: gap of unknown length
* 1103 2301: contig of 1199 bp in length
* 2302 2401: gap of unknown length
* 2402 3505: contig of 1104 bp in length

```

* 3506 3605: gap of unknown length
* 3606 4783: contig of 1178 bp in length
* 4784 4883: gap of unknown length
* 4884 6056: contig of 1173 bp in length
* 6057 6156: gap of unknown length
* 6157 7495: contig of 1339 bp in length
* 7496 7595: gap of unknown length
* 7596 8658: contig of 1063 bp in length
* 8659 8758: gap of unknown length
* 8759 10238: contig of 1480 bp in length
* 10239 10338: gap of unknown length
* 10339 12220: contig of 1882 bp in length
* 12221 12320: gap of unknown length
* 12321 13477: contig of 1157 bp in length
* 13478 13577: gap of unknown length
* 13578 14605: contig of 1028 bp in length
* 14606 14705: gap of unknown length
* 14706 15755: contig of 1050 bp in length
* 15756 15855: gap of unknown length
* 15856 17702: contig of 1847 bp in length
* 17703 17802: gap of unknown length
* 17803 19472: contig of 1670 bp in length
* 19473 19572: gap of unknown length
* 19573 21485: contig of 1913 bp in length
* 21486 21585: gap of unknown length
* 21586 23432: contig of 1847 bp in length
* 23433 23532: gap of unknown length
* 23533 25029: contig of 1497 bp in length
* 25030 25129: gap of unknown length
* 25130 26659: contig of 1530 bp in length
* 26660 26759: gap of unknown length
* 26760 28350: contig of 1591 bp in length
* 28351 28450: gap of unknown length
* 28451 30077: contig of 1627 bp in length
* 30078 30177: gap of unknown length
* 30178 31793: contig of 1616 bp in length
* 31794 31893: gap of unknown length
* 31894 33673: contig of 1780 bp in length
* 33674 33773: gap of unknown length
* 33774 34989: contig of 1216 bp in length
* 34990 35089: gap of unknown length
* 35090 36437: contig of 1348 bp in length
* 36438 36537: gap of unknown length
* 36538 37948: contig of 1411 bp in length
* 37949 38048: gap of unknown length
* 38049 39244: contig of 1196 bp in length
* 39245 39344: gap of unknown length
* 39345 41409: contig of 2065 bp in length
* 41410 41509: gap of unknown length
* 41510 43089: contig of 1580 bp in length
* 43090 43189: gap of unknown length
* 43190 45987: contig of 2798 bp in length
* 45988 46087: gap of unknown length
* 46088 47997: contig of 1910 bp in length
* 47998 48097: gap of unknown length
* 48098 50569: contig of 2472 bp in length
* 50570 50669: gap of unknown length
* 50670 53411: contig of 2742 bp in length
* 53412 53511: gap of unknown length
* 53512 56398: contig of 2887 bp in length
* 56399 56498: gap of unknown length
* 56499 58805: contig of 2307 bp in length
* 58806 58905: gap of unknown length
* 58906 61202: contig of 2297 bp in length
* 61203 61302: gap of unknown length
* 61303 63871: contig of 2569 bp in length
* 63872 63971: gap of unknown length
* 63972 66802: contig of 2831 bp in length
* 66803 66902: gap of unknown length
* 66903 69598: contig of 2696 bp in length
* 69599 72577: gap of unknown length
* 72578 72677: contig of 2879 bp in length

```

```

Query Match          92.0%; Score 18.4; DB 1; Length 164552;
Best Local Similarity 95.0%; Pred. No. 19;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 TTAGTGTAGTATTAGTT 20
Db 130778 TTAGTGTAGTATTAGTT 130797

```

```

* 72678 75917: contig of 3240 bp in length
* 75918 76017: gap of unknown length
* 76018 79500: contig of 3483 bp in length
* 79501 79600: gap of unknown length
* 79601 81781: contig of 2181 bp in length
* 81782 81881: gap of unknown length
* 81882 84868: contig of 2987 bp in length
* 84869 84968: gap of unknown length
* 84969 89049: contig of 4081 bp in length
* 89050 89149: gap of unknown length
* 89150 94104: contig of 4955 bp in length
* 94105 94204: gap of unknown length
* 94205 97324: contig of 3120 bp in length
* 97325 97424: gap of unknown length
* 97425 102026: contig of 4602 bp in length
* 102027 102126: gap of unknown length
* 102127 106368: contig of 4242 bp in length
* 106369 106468: gap of unknown length
* 106469 111386: contig of 4918 bp in length
* 111387 111486: gap of unknown length
* 111487 117245: contig of 5759 bp in length
* 117246 117345: gap of unknown length
* 117346 121639: contig of 4294 bp in length
* 121640 121739: gap of unknown length
* 121740 125932: contig of 4193 bp in length
* 125933 126032: gap of unknown length
* 126033 132259: contig of 6227 bp in length
* 132260 132359: gap of unknown length
* 132360 137015: contig of 4636 bp in length

```

```

RESULT 11
LOCUS          AC113369          176400 bp    DNA     linear   HTG 01-MAR-2002
DEFINITION    Homo sapiens chromosome 5 clone RP11-119P10, WORKING DRAFT
ACCESSION     AC113369
VERSION       AC113369.1  GI:19033483
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVERIN.
SOURCE        Homo sapiens.
ORGANISM      Homo sapiens
REFERENCE     1 (bases 1 to 176400)
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE         Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
JOURNAL       Submitted (01-MAR-2002) Production Sequencing Facility, DOE Joint
              Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT      -----Genome Center
              Center: Joint Genome Institute
              Center Code: JGI
              Web site: http://www.jgi.doe.gov

```

```

Project Information
Center Project Name: 452431
Center Clone name: RPCI-11_119P10
-----
Summary Statistics
Consensus quality: 175634 bases at least Q40
Consensus quality: 176221 bases at least Q30
Consensus quality: 176375 bases at least Q20

```

Estimated insert size: 182000; agarose-fp estimation
 Estimated insert size: 176400; sum-of-contigs estimation
 Quality coverage: 12.49 in Q20 bases; agarose-fp estimation
 Quality coverage: 12.88 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 176400: contig of 176400 bp in length.
 Location/Qualifiers
 1. 176400
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="RP11-119P10"
 /clone_id="RPCT human BAC library 11"
 BASE COUNT 55190 a 37584 c 34956 g 48670 t
 ORIGIN

Query Match 92.0% Score 18.4: DB 1: length 176400:
 Best Local Similarity 95.0% Pred. No. 20;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 TTACTGTGATGTTATGTT 20
 ||||| ||||| ||||| |||||
 Db 78311 TTACTGTGATGTTATGTT 78330

RESULT 12
 AC115176/c
 LOCUS
 DEFINITION Rattus norvegicus clone CH230-138122, *** SEQUENCING IN PROGRESS
 AC115176 178712 bp DNA linear HTG 13-JUL-2002
 *** 63 unordered pieces.
 AC115176
 VERSION AC115176.5 GI:21738734
 KEYWORDS HTG: HTGS_PHASE1.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 178712)
 Musny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barberia,J., Benton,J., Bimage,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhaý,C., Burch,P., Burkelt,C., Burrell,K.L., Byrd,N.C.,
 Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falla,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabali,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hognes,M., Holloway,C., Hollins,B.,
 Homs,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratochic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louised,H.,
 Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapa,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogih,M., Okunou,G.,
 Orangunye,N., Oyiedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rivers,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Severy,G.,
 Scherer,S., Scott,G., Shen,H., Shoshitari,N., Stison,I.,
 Sodergren,E., Sonaike,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
 Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstock,G., and Gibbs,R.
 Direct Submission
 Unpublished
 2 (bases 1 to 178712)
 Worley,K.C.
 Direct Submission
 Submitted (15-MAR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 178712)
 Worley,K.C.
 Direct Submission
 Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Jul 12, 2002 this sequence version replaced gi:20467608.
 ----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: http://www.hgsc.bcm.tmc.edu/
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GRG5
 Center clone name: CH230-138122
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 121130 bases at least Q40
 Consensus quality: 127396 bases at least Q30
 Consensus quality: 132791 bases at least Q20

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 63 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1025: contig of 1025 bp in length
 1026 1125: gap of unknown length
 1126 2503: contig of 1378 bp in length
 2504 2603: gap of unknown length
 2604 3647: contig of 1044 bp in length
 3648 3747: gap of unknown length
 3748 4848: contig of 1101 bp in length
 4849 4948: gap of unknown length
 4949 6516: contig of 1568 bp in length
 6517 7633: gap of unknown length
 7634 7733: contig of 1017 bp in length
 7734 7734: gap of unknown length
 7734 9412: contig of 1679 bp in length
 9413 9512: gap of unknown length
 9513 10564: contig of 1052 bp in length
 10565 10665: gap of unknown length
 10666 11733: contig of 1069 bp in length
 11734 11833: gap of unknown length
 11834 13300: contig of 1467 bp in length

13301 13400: gap of unknown length
13401 15196: contig of 1796 bp in length
15197 15296: gap of unknown length
15297 16682: contig of 1386 bp in length
16683 16782: gap of unknown length
16783 17945: contig of 1163 bp in length
17946 18045: gap of unknown length
18046 19776: contig of 1731 bp in length
19777 19876: gap of unknown length
19877 21957: contig of 2081 bp in length
21958 22057: gap of unknown length
22058 23188: contig of 1131 bp in length
23189 24464: contig of 1176 bp in length
24465 24564: gap of unknown length
24565 25908: contig of 1344 bp in length
25909 26008: gap of unknown length
26009 27814: contig of 1806 bp in length
27815 27914: gap of unknown length
27915 29387: contig of 1473 bp in length
29388 29487: gap of unknown length
29488 30801: contig of 1314 bp in length
30802 30901: gap of unknown length
30902 32199: contig of 1298 bp in length
32200 32299: gap of unknown length
32300 33740: contig of 1441 bp in length
33741 33840: gap of unknown length
33841 34992: contig of 1152 bp in length
34993 35092: gap of unknown length
35093 37835: contig of 2743 bp in length
37836 37935: gap of unknown length
37936 41035: contig of 3100 bp in length
41036 41135: gap of unknown length
41136 43185: contig of 2050 bp in length
43186 43285: gap of unknown length
43286 45335: contig of 2050 bp in length
45336 45435: gap of unknown length
45436 46618: contig of 1183 bp in length
46619 46719: gap of unknown length
46719 49009: contig of 2291 bp in length
49010 49109: gap of unknown length
49110 52162: contig of 3053 bp in length
52163 52262: gap of unknown length
52263 54282: contig of 2020 bp in length
54283 54382: gap of unknown length
54383 56291: contig of 1909 bp in length
56292 56391: gap of unknown length
56392 58037: contig of 1646 bp in length
58038 58137: gap of unknown length
58138 62042: contig of 3905 bp in length
62043 62142: gap of unknown length
62143 64769: contig of 2627 bp in length
64770 64870: gap of unknown length
64870 67320: contig of 2451 bp in length
67321 67420: gap of unknown length
67421 70041: contig of 2621 bp in length
70042 70141: gap of unknown length
70142 73105: contig of 2964 bp in length
73106 73205: gap of unknown length
73206 76419: contig of 3214 bp in length
76420 76519: gap of unknown length
76520 79593: contig of 3074 bp in length
79594 79693: gap of unknown length
79694 82683: contig of 2990 bp in length
82684 82783: gap of unknown length
82784 85451: contig of 2668 bp in length
85452 85551: gap of unknown length
85552 89457: contig of 3906 bp in length
89458 89557: gap of unknown length
89558 91719: contig of 2162 bp in length
91720 91819: gap of unknown length
91820 95602: contig of 3783 bp in length
95603 95702: gap of unknown length

Query Match 92.0%: Score 18.4; DB 1; Length 178712;
Best Local Similarity 95.0%: Pred. No. 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTAGTGTAGTTAGTTAGTT 20
Db 109564 TTAGTGTAGTTAGTTAGTT 109545

RESULT 13
AC012589
LOCUS
DEFINITION Homo sapiens chromosome 5 clone RP11-119P10 map 5, WORKING DRAFT
SEQUENCE, 23 unordered pieces.
ACCESSION AC012589
VERSION AC012589.5 GI:7387335
KEYWORDS HTG: HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eumetazoa; Euthera; Primates; Carnivora; Mammalia; Homo.
TITLE 1 (bases 1 to 178970)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
JOURNAL Homo sapiens chromosome 5, clone RP11-119P10
REFERENCE
AUTHORS 2 (bases 1 to 178970)
JOURNAL Unpublished

TITLE Direct Submission
JOURNAL Submitted (30-Oct-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 178970)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Boguslavsky, L., Boukhvalter, B., Brown, A.,
Camarata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,
Collins, S., Collymore, A., Cooke, F., DeRellano, K., Dewar, K.,
Diaz, J., S., Dodge, S., Faro, S., Ferrelita, P., Fitzhugh, W., Gage, D.,
Galagan, J., Gardyna, S., Glinde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Huine, N.,
Iliev, I., Johnson, R., Jones, C., Karatas, A., Labocque, K.,
Lamarez, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., MacDonald, P., Marquis, N., Matthews, C., McCarthy, M.,
McEwan, P., McKernan, K., McPherson, R., Melidim, J., Mennas, L.,
Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Sudmanian, A., Talamas, J.,
Tessaye, S., Threll, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyma, D., Ye, W.J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Peterson, K., Phunhkhang, P., Pierre, N., Pollara, V., Raymond, C.,
Retta, R., Rieback, M., Riley, R., Riese, C., Rogov, P., Roman, D.,
Roselli, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S.,
Severy, P., Sounez, C., Spencer, B., Stange-Thomann, N.,
Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Testaye, S.,
Theodore, J., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 3, 2000 this sequence version replaced gi:6910762.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: MIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L3738

Center clone name: 119_P.10

----- Summary Statistics

Sequencing vector: M13; M7815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 160436 bases at least Q40
Consensus quality: 168973 bases at least Q30
Consensus quality: 173202 bases at least Q20
Insert size: 177000; agarose-fp
Insert size: 176770; sum-of-coverage
Quality coverage: 4.0 in Q20 bases; sum-of-coverage

* NOTE: This is a 'working draft' sequence. It currently
* consists of 23 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
* 1012 1111: contig of 1011 bp in length
* 1112 2722: contig of 1611 bp in length
* 2723 2822: gap of 100 bp
* 2823 4300: contig of 1478 bp in length
* 4301 4400: gap of 100 bp
* 4401 5425: contig of 1025 bp in length
* 5426 5525: gap of 100 bp
* 5526 7534: contig of 2009 bp in length
* 7535 7634: gap of 100 bp
* 7635 9780: contig of 2146 bp in length
* 9781 9880: gap of 100 bp
* 9881 12462: contig of 2582 bp in length
* 12463 12562: gap of 100 bp
* 12563 15626: contig of 3064 bp in length
* 15627 15726: gap of 100 bp
* 15727 17635: contig of 1909 bp in length
* 17636 17735: gap of 100 bp
* 17736 19894: contig of 2159 bp in length
* 19895 19994: gap of 100 bp
* 19995 25298: contig of 5304 bp in length
* 25299 25398: gap of 100 bp
* 25399 30406: contig of 5008 bp in length
* 30407 30506: gap of 100 bp
* 30507 38666: contig of 8160 bp in length
* 38667 38766: gap of 100 bp
* 38767 45449: contig of 6683 bp in length
* 45450 45549: gap of 100 bp
* 45550 55180: contig of 9631 bp in length
* 55181 55280: gap of 100 bp

FEATURES

source

* 55281 63752: contig of 8472 bp in length
* 63753 63852: gap of 100 bp
* 63853 75295: contig of 11443 bp in length
* 75296 75395: gap of 100 bp
* 75396 87583: contig of 12188 bp in length
* 87584 87683: gap of 100 bp
* 87684 98245: contig of 10562 bp in length
* 98246 98345: gap of 100 bp
* 98346 11685: contig of 13340 bp in length
* 11686 11785: gap of 100 bp
* 11786 131002: contig of 19217 bp in length
* 131003 131103: gap of 100 bp
* 131103 152070: contig of 20968 bp in length
* 152071 152170: gap of 100 bp
* 152171 178970: contig of 26800 bp in length.
location/Qualifiers
1. 178970
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="5"
/map="5"
/clone="RP11-119P10"
/clone_lib="RP11 Human Male BAC"
1. 1011
/note="assembly-fragment"
1112. 2722
/note="assembly-fragment"
2823. 4300
/note="assembly-fragment"
4401. 5425
/note="assembly-fragment"
5526. 7534
/note="assembly-fragment"
7635. 9780
/note="assembly-fragment"
9881. 12462
/note="assembly-fragment"
12563. 15626
/note="assembly-fragment"
15727. 17635
/note="assembly-fragment"
17736. 19894
/note="assembly-fragment"
19995. 25298
/note="assembly-fragment"
25399. 30406
/note="assembly-fragment"
30507. 38666
/note="assembly-fragment"
38767. 45449
/note="assembly-fragment"
45550. 55180
/note="assembly-fragment"
55281. 63752
/note="assembly-fragment"
63853. 75295
/note="assembly-fragment"
75396. 87583
/note="assembly-fragment"
87684. 98245
/note="assembly-fragment"
98346. 111685
/note="assembly-fragment"
111786. 131002
/note="assembly-fragment"
131103. 152070
/note="assembly-fragment"
clone_end:SP6
vector_side:right
152171. 178970
/note="assembly-fragment"
clone_end:T7
vector_side:left"


```

BASE COUNT      54765 a 37234 c 35721 g 49050 t 2200 others
ORIGIN
Query Match      92.0%: Score 18.4; DB 1; Length 178970;
Best Local Similarity 95.0%: Pred. No. 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 TTAGTGTAGTATTAGTT 20
      ||||| ||||| ||||| |||||
Db 22509 TTAGTATTAGTTATTAGTT 22528

RESULT 14
AC122335      201132 bp   DNA      linear   HTG 02-AUG-2002
LOCUS      Mus musculus chromosome UNK clone RP23-350F7, WORKING DRAFT
DEFINITION
AC122335
AC122335.2 GI:22091367
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE      1 (bases 1 to 201132)
JOURNAL      The sequence of Mus musculus clone
REFERENCE
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE      2 (bases 1 to 201132)
JOURNAL      Unpublished
REFERENCE
AUTHORS      McPherson,J.D. and Waterston,R.H.
TITLE      Direct Submission
JOURNAL      Submitted (23-MAY-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 201132)
McPherson,J.D. and Waterston,R.H.
Direct Submission
JOURNAL      Submitted (02-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
On Aug 2, 2002 this sequence version replaced gi:21105196.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_BA0350F07
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: Plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; Version 0.990319
Consensus quality: 0 bases at least Q40
Consensus quality: 0 bases at least Q30
Consensus quality: 0 bases at least Q20
Insert size: 191000; agarose-fp
Insert size: 199573; sum-of-contigs
Quality coverage: 14.48 in Q20 bases; sum-of-contigs
Quality coverage: 11.45 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2374: contig of 2374 bp in length
* 2375 2474: gap of unknown length
* 2475 7198: contig of 4724 bp in length

```

```

* 7199 7298: gap of unknown length
* 7299 13971: contig of 6673 bp in length
* 13972 14071: gap of unknown length
* 14072 23201: contig of 9130 bp in length
* 23202 23301: gap of unknown length
* 23302 36024: contig of 12723 bp in length
* 36025 36125: gap of unknown length
* 36125 55431: contig of 19306 bp in length
* 55431 55531: gap of unknown length
* 55531 75591: contig of 20060 bp in length
* 75591 75691: gap of unknown length
* 75691 100853: contig of 25163 bp in length
* 100854 100953: gap of unknown length
* 100954 200373: contig of 99420 bp in length
* 200374 200473: gap of unknown length
* 200474 201132: contig of 659 bp in length.
FEATURES
source
1..201132
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="UNK"
/clone="RP23-350F7"
1..2374
/feature="assembly_name:Contig11"
2475..7198
/feature="assembly_name:Contig12"
7299..13971
/feature="assembly_name:Contig13"
14072..23201
/feature="assembly_name:Contig14"
23302..36024
/feature="assembly_name:Contig15"
36125..55430
/feature="assembly_name:Contig16"
55531..75590
/feature="assembly_name:Contig17"
75691..100853
/feature="assembly_name:Contig18"
100954..200373
/feature="assembly_name:Contig19"
200474..201132
/feature="assembly_name:Contig9"
BASE COUNT      56228 a 41046 c 42038 g 60912 t 908 others
ORIGIN
Query Match      92.0%: Score 18.4; DB 1; Length 201132;
Best Local Similarity 95.0%: Pred. No. 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 1 TTAGTGTAGTATTAGTT 20
      ||||| ||||| ||||| |||||
Db 143770 TTAGTGTAGTATTAGTT 143789

RESULT 15
AL844536      241585 bp   DNA      linear   HTG 09-AUG-2002
LOCUS      Mus musculus chromosome 2 clone RP23-22A15, *** SEQUENCING IN
PROGRESS ***; 11 unordered pieces.
DEFINITION
AL844536
AL844536.3 GI:22204804
KEYWORDS      HTG; HTGS_PHASE1.
SOURCE      house mouse.
ORGANISM      Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS      Plumb,B.
TITLE      1 (bases 1 to 241585)
JOURNAL      Submitted (02-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
hunger@anger.ac.uk Clone requests: clonerequestsanger.ac.uk
On Aug 11, 2002 this sequence version replaced gi:21955751.
COMMENT

```

```

/note="assembly_fragment:01569
fragment_chain:2"
misc.feature
102018..146947
/note="assembly_fragment:00752
fragment_chain:2
clone_end:T7
vector_side:right"
147048..187542
/note="assembly_fragment:00109
fragment_chain:2"
187643..192747
/note="assembly_fragment:00452
fragment_chain:2"
192848..241585
/note="assembly_fragment:00041
fragment_chain:2"
misc.feature
BASE COUNT      68487 a 53697 c 52331 g 66069 t 1001 others
ORIGIN
Query Match      92.0%; Score 18.4; DB 1; Length 241585;
Best Local Similarity 95.0%; Pred. No. 20;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY      1 TTAGTGTACTTATTAGTT 20
||||| |||||||||
Db 164303 TTAGTTATTAGTTATTAGTT 164322

Search completed: April 6, 2003, 11:56:21
Job time : 524.5 secs

```

GenCore version 5.1.4.p5_A578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 10:34:07 ; Search time 243 Seconds
(without alignments)
1069.066 Million cell updates/sec

Title: US-09-980-265-12

Perfect score: 20

Sequence: 1 ttagtltgtagttagt 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1994485 seqs, 6494577260 residues

Total number of hits satisfying chosen parameters: 3988970

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl_NOHTG:*

1: gb_ba:*
2: gb_in:*
3: gb_om:*
4: gb_ov:*
5: gb_pat:*
6: gb_ph:*
7: gb_pl:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vl:*
14: em_ba:*
15: em_fun:*
16: em_hum:*
17: em_in:*
18: em_mu:*
19: em_om:*
20: em_or:*
21: em_ov:*
22: em_pat:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_vl:*
29: em_htg_hum:*
30: em_htg_inv:*
31: em_htg_other:*
32: em_htg_mus:*
33: em_htg_pln:*
34: em_htg_rod:*
35: em_htg_mam:*
36: em_htg_vrt:*
37: em_sy:*
38: em_htgo_hum:*
39: em_htgo_mus:*
40: em_htgo_other:*

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	20	100.0	20	5	AX057378	AX057378 Sequence
2	18.4	92.0	20	5	AX057367	AX057367 Sequence
3	18.4	92.0	20	5	AX057371	AX057371 Sequence
4	18.4	92.0	20	5	AX057375	AX057375 Sequence
5	18.4	92.0	20	5	AX057380	AX057380 Sequence
6	18.4	92.0	20	5	AX057381	AX057381 Sequence
7	18.4	92.0	20	5	AX057382	AX057382 Sequence
8	18.4	92.0	2658	1	AF108143	AF108143 Anabaena
9	18.4	92.0	2876	1	ASU08013	U08013 Anabaena sp
10	18.4	92.0	3568	1	AF038558	AF038558 Fischerel
11	18.4	92.0	11029	5	AX251168	AX251168 Sequence
12	18.4	92.0	16258	5	AX348425	AX348425 Sequence
13	18.4	92.0	123078	7	AX348808	AX348808 Sequence
14	18.4	92.0	165820	8	AC010232	AP004902 Lotus jap
15	18.4	92.0	174919	2	AC009370	AC010232 Homo sapi
16	18.4	92.0	175867	2	AC007976	AC009370 Drosophill
17	18.4	92.0	178628	2	AC009385	AC007976 Drosophill
18	18.4	92.0	292326	2	AE003520	AC009385 Drosophill
19	18.4	92.0	303250	1	AP003193	AE003520 Drosophill
20	18.4	92.0	324484	2	AE003584	AP003193 Clostridi
21	18.4	92.0	339650	1	AP003583	AE003584 Drosophill
22	18.4	92.0	347550	1	AP003595	AP003583 Nostoc sp
23	18.4	92.0	349050	1	AP003586	AP003595 Nostoc sp
24	18.4	90.0	177402	8	AC007406	AP003586 Nostoc sp
25	17.4	87.0	5293	1	AX346560	AC007406 Homo sapi
26	17.4	87.0	7238	5	AX345350	L14265 Leptospira
27	17.4	87.0	9507	5	AX344835	AX346560 Sequence
28	17.4	87.0	9728	5	AX346805	AX345350 Sequence
29	17.4	87.0	15698	5	AX348479	AX344835 Sequence
30	17.4	87.0	15698	5	AX347043	AX348479 Sequence
31	17.4	87.0	15698	5	AX345427	AX347043 Sequence
32	17.4	87.0	15698	5	AX345427	AX345427 Sequence
33	17.4	87.0	15698	5	AX345427	AX345427 Sequence
34	17.4	87.0	15698	5	AX345427	AX345427 Sequence
35	17.4	87.0	15698	5	AX345427	AX345427 Sequence
36	17.4	87.0	15698	5	AX345427	AX345427 Sequence
37	17.4	87.0	15698	5	AX345427	AX345427 Sequence
38	17.4	87.0	15698	5	AX345427	AX345427 Sequence
39	17.4	87.0	15698	5	AX345427	AX345427 Sequence
40	17.4	87.0	15698	5	AX345427	AX345427 Sequence

ALIGNMENTS

RESULT 1
AX057378
LOCUS AX057378 20 bp DNA
DEFINITION Sequence 12 from Patent WO0075304.
ACCESSION AX057378
VERSION AX057378.1 GI:12310119
KEYWORDS
SOURCE
ORGANISM
SYNTHETIC CONSTRUCT.
ARTIFICIAL SEQUENCES.
REFERENCE
1 (bases 1 to 20)
Bachy M., Sodoyer R. and Tranomy E.
Immunostimulant oligonucleotide
Patent: WO 0075304-A 12 14-DEC-2000;
JOURNAL
Aventis Pasteur (FR)
Location/Qualifiers
FEATURES

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,

```
source 1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"
BASE COUNT 4 a 0 c 4 g 12 t
ORIGIN

Query Match 100.0%; Score 20; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TTAGTTGTTAGTTATTAGTT 20
Db 1 TTAGTTGTTAGTTATTAGTT 20

RESULT 2
AX057367 20 bp DNA linear PAT 17-JAN-2001
LOCUS
DEFINITION Sequence 1 from Patent WO0075304.
ACCESSION AX057367
VERSION AX057367.1 GI:12310108
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 1 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"
BASE COUNT 5 a 0 c 3 g 12 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTTGTTAGTTATTAGTT 20
Db 1 TTAGTTGTTAGTTATTAGTT 20

RESULT 3
AX057371 20 bp DNA linear PAT 17-JAN-2001
LOCUS
DEFINITION Sequence 5 from Patent WO0075304.
ACCESSION AX057371
VERSION AX057371.1 GI:12310112
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 5 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"
BASE COUNT 4 a 0 c 3 g 13 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTTGTTAGTTATTAGTT 20
Db 1 TTAGTTGTTAGTTATTAGTT 20

RESULT 4
AX057375 20 bp DNA linear PAT 17-JAN-2001
LOCUS
DEFINITION Sequence 9 from Patent WO0075304.
ACCESSION AX057375
VERSION AX057375.1 GI:12310116
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 9 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"
BASE COUNT 4 a 1 c 3 g 12 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTTGTTAGTTATTAGTT 20
Db 1 TTAGTTGTTAGTTATTAGTT 20

RESULT 5
AX057380 20 bp DNA linear PAT 17-JAN-2001
LOCUS
DEFINITION Sequence 14 from Patent WO0075304.
ACCESSION AX057380
VERSION AX057380.1 GI:12310121
KEYWORDS
SOURCE synthetic construct.
ORGANISM artificial sequences.
REFERENCE 1 (bases 1 to 20)
AUTHORS Bachy,M., Sodoyer,R. and Tranoy,E.
TITLE Immunostimulant oligonucleotide
JOURNAL Patent: WO 0075304-A 14 14-DEC-2000;
Aventis Pasteur (FR)
FEATURES
source location/Qualifiers
1..20
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="oligonucleotide"
BASE COUNT 3 a 0 c 4 g 13 t
ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 TTAGTTGTTAGTTATTAGTT 20
Db 1 TTAGTTGTTAGTTATTAGTT 20

RESULT 6
AX057381
```

```
LOCUS       AX057381                20 bp    DNA
DEFINITION  Sequence 15 from Patent W00075304.
ACCESSION   AX057381
VERSION     AX057381.1  GI:12310122
KEYWORDS    .
SOURCE      synthetic construct.
ORGANISM    synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 20)
AUTHORS     Bachy,M., Sodoyer,R. and Trannoy,E.
TITLE       Immunostimulant oligonucleotide
            Patent: WO 0075304-A 15 14-DEC-2000;
            Aventis Pasteur (FR)
JOURNAL     .
FEATURES    source
            1..20
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="Oligonucleotide"
BASE COUNT  3 a                1 c                4 g                12 t
ORIGIN
Query Match      92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TTAGTTGTTAGTTATTAGTT 20
        |||
Db      1 TTAGTTGTTAGTTCTTACTT 20

RESULT 7
AX057382                20 bp    DNA
DEFINITION  Sequence 16 from Patent W00075304.
ACCESSION   AX057382
VERSION     AX057382.1  GI:12310123
KEYWORDS    .
SOURCE      synthetic construct.
ORGANISM    synthetic construct
            artificial sequences.
REFERENCE   1 (bases 1 to 20)
AUTHORS     Bachy,M., Sodoyer,R. and Trannoy,E.
TITLE       Immunostimulant oligonucleotide
            Patent: WO 0075304-A 16 14-DEC-2000;
            Aventis Pasteur (FR)
JOURNAL     .
FEATURES    source
            1..20
            /organism="synthetic construct"
            /db_xref="taxon:32630"
            /note="Oligonucleotide"
BASE COUNT  3 a                0 c                5 g                12 t
ORIGIN
Query Match      92.0%; Score 18.4; DB 5; Length 20;
Best Local Similarity 95.0%; Pred. No. 1.4e+03;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TTAGTTGTTAGTTATTAGTT 20
        |||
Db      1 TTAGTTGTTAGTTGTTAGTT 20

RESULT 8
AF108143/c                2658 bp    DNA
DEFINITION  Anabaena PCC7120 Ni-Fe uptake hydrogenase small and large subunit
            genes, upstream sequence and unknown gene.
ACCESSION   AF108143
VERSION     AF108143.1  GI:5353551
KEYWORDS    .
SOURCE      Nostoc sp. PCC 7120.
            Nostoc sp. PCC 7120
            Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.

REFERENCE   1 (bases 1 to 2658)
AUTHORS     Golden,J.W.
TITLE       Anabaena sp. genomic sequence upstream of [nife] uptake hydrogenase
            small subunit (nups) gene and large subunit (nupL) gene
JOURNAL     unpublished
REFERENCE   2 (bases 1 to 2658)
AUTHORS     Golden,J.W.
TITLE       Direct Submission
JOURNAL     Submitted (23-NOV-1998) Department of Biology, Texas A&M
            University, College Station, TX 77843-3258, USA
FEATURES    source
            1..2658
            /organism="Nostoc sp., PCC 7120"
            /db_xref="taxon:103690"
            /clone="pAM1268"
            /note="region upstream of Ni-Fe uptake hydrogenase small
            and large subunit genes"
            complement(510..1499)
            /codon_start=1
            /transl_table=1
            /product="unknown"
            /protein_id="AAD42178.1"
            /db_xref="GI:5353552"
            /translation="MTPSITDSTYKAAALAAISSEATTTEKIQMLIELAQIQKPKPT
            PGDLMAVELVOQAIKKCGDYSIMKRSQGMAGALSKSIDPGAEILLQKAGYEBA
            LPVRLAAAEVAEAGQMGFLVLOSILPEFLARITDSIAAYOGAMGVTSODYPOEY
            AILANNVAIATLSMSGSGEONLFAGLAVOCTPEYATANKKNAALFNLDPDEKPE
            LQYLPSPHPENNRIATATIEALIKVRCOCTPEYATANKKNAALFNLDPDEKPE
            LGNSQNLQAAVAYVOEAMEIFSQCOEQAQVAQALAEVAAELIRISMSNGDK"
BASE COUNT  838 a                664 c                423 g                733 t
ORIGIN
Query Match      92.0%; Score 18.4; DB 1; Length 2658;
Best Local Similarity 95.0%; Pred. No. 2.9e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 TTAGTTGTTAGTTATTAGTT 20
        |||
Db      2658 TTAGTTGTTAGTTGTTAGTT 2639

RESULT 9
ASU08013/c                2876 bp    DNA
DEFINITION  Anabaena sp. [Nife] uptake hydrogenase small subunit (nups) and
            [Nife] uptake hydrogenase large subunit (nupL) genes, complete cds.
ACCESSION   ASU08013
VERSION     ASU08013.1  GI:3929403
KEYWORDS    .
SOURCE      Anabaena sp.
ORGANISM    Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Anabaena.
REFERENCE   1 (bases 918 to 2876)
AUTHORS     Carrasco,C.D., Buettner,J.A. and Golden,J.W.
TITLE       Programmed DNA rearrangement of a cyanobacterial nupL gene in
            heterocysts
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 92 (3), 791-795 (1995)
MEDLINE    95148623
PUBMED     7846053
REFERENCE   2 (bases 1 to 2876)
AUTHORS     Carrasco,C.D., Holliday,S., Garcia,J.S. and Golden,J.W.
TITLE       xisc encodes a site-specific recombinase required for the
            developmentally regulated excision of the nupL element in Anabaena
            unpublished
JOURNAL     3 (bases 918 to 2876)
            Golden,J.W.
            Direct Submission
            Submitted (24-MAR-1994) Biology, Texas A&M University, College
            Station, TX 77843-3258, USA
REFERENCE   4 (bases 1 to 2876)
AUTHORS     Carrasco,C.D. and Golden,J.W.
TITLE       Direct Submission
JOURNAL     Submitted (30-NOV-1998) Biology, Texas A&M University, College
```

Station, TX 77843-3258, USA
REMARK Sequence updated by submitter
COMMENT On Nov 30, 1998 this sequence version replaced gi:520458.
FEATURES
source
/organism="Anabaena sp."
/strain="PCC 7120"
/db_xref="taxon:1167"
/map="On Bln1-E and Psl1-A between 0.4 and 0.9 Mb"
/clone="PAM1268, PAM1314"
/cell_type="heterocyst"
96..1072
/gene="hups"
96..1058
/gene="hups"
/codon_start=1
/transl_table=11
/product="[NifE] uptake hydrogenase small subunit"
/protein_id="AAC79877.1"
/db_xref="GI:3929404"
/translation="MTNVLMLGAGSGNTMSFLNAEPTVCDLIADFGIKVLMHPSL
GLEIDNLTLDICILKIPDLIVFEGVYVNAPEGTGEMNFADRPKMDLITDLAG
AASFVAVGDCATWGGIIPAMEPNSGLOFLKRGGGFLGODEFTKSGIPYINIG
CAHPDWITQILVAITGRIGDLADELNRPQTFNTPTOTGCTRNVRHAKATTAEP
GOKKGLFTDLGCGPMTSSCNRIILMNVSSTRAGMPCLGCTEPEFPFDLKPGTV
FKQTITGVPEKELPGVSNKNYAVLTMVAKDAPKVAEDDEFTV"
1127..2722
/gene="hupL"
1127..2722
/gene="hupL"
/note="rearranged hupL gene on heterocyst chromosome"
/codon_start=1
/transl_table=11
/product="[NifE] uptake hydrogenase large subunit"
/protein_id="AAC79878.1"
/db_xref="GI:520460"
/translation="MTIKTLDISPVGRVSGDLVREIEDGRVNAWTHAEFRGEFI
ILNGKPOAGLIVTPRICIGASHLTASMAIDTAMNTVPNNALIANLQIVETI
OSIPRYEGLFALDITNKRSRFRSAYTAKSYELGVTISLKEPEIYALF
GGOMPHSSVPGVNCAPLITDITRAMALIEYFTNMLEPVMILGSLREYEIOTYD
DMDWLEADIKHRESDLGFWRMGLDIGLRGAGVGYKVSNGVILPHEKXKOPTIEG
RNAAMTKSGVDSFENTHTLMDHTFARENTTHAWYDEGNADHPDRTTKPRHKKTK
DERNAISWSTAVLHODFGRLEVEGPLAROLVAGGQHCESWQHYDFTLDLQKMGASIS
HLROLARVHEIVKLYROARCLREFVLANDPWYIKPEKGRGATGATASGSLCHMID
IEGGKIKNOYVIAATITMNVGPRDSEGVGRPIEBALIGTIEDSRDEVEGVHARSFDS
CVCYVHAHDARKGAEELARFRA"
1622..1637
/gene="hupL"
/note="hupL element recombination site on heterocyst
chromosome"
BASE COUNT 777 a 728 c 670 g 701 t
ORIGIN
Query Match 92.0%; Score 18.4; DB 1; Length 2876;
Best Local Similarity 95.0%; Pred. No. 2.8e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTAGTCTAGTATATAGTT 20
||||| 1111111111111111
DB 39 TTAGTCTAGTCTAGTT 20

REFERENCE 1 (bases 1 to 5568)
AUTHORS Sun,J., He,Z.Y., Nechushtai,R. and Chitnis,P.R.
TITLE Molecular cloning of the psaa and psab genes for the core proteins
of photosystem I from the thermophilic cyanobacterium Mastigocladus
lamninus (Accession No. AF038558) (PCR98-041)
JOURNAL Plant Physiol. 116 (3), 1192 (1998)
REFERENCE 2 (bases 1 to 5568)
AUTHORS Sun,J., He,Z.Y., Nechushtai,R. and Chitnis,P.R.
TITLE Direct Substitution
JOURNAL Submitted (13-DEC-1997) Biochemistry and Biophysics, Iowa State
University, 4156 Mol. Biol. Bldg., Ames, IA 50011, USA
FEATURES
source
Location/Qualifiers
1..5568
/organism="Fischerella sp. PCC 7605"
/strain="PCC7605"
/db_xref="taxon:60455"
/note="submitted as Mastigocladus lamninus"
593..2851
/gene="psaa"
593..2851
/gene="psaa"
/note="psaa"
/gene="psaa"
/note="binds chlorophylls, iron-sulfur centers and
phyllonones"
/codon_start=1
/transl_table=11
/product="photosystem I psaa"
/protein_id="AAC12866.1"
/db_xref="GI:2827200"
/translation="MTISPPEREKKARYIVNDPVPYSEELMAKPGHFDRLTARPK
TTTWIYNLALAHDPDTHSDLEDSRKIFAAHFGILAVNTLMISMTIRGARSYE
AMISDLVHVPKSAQVAVPIVGDITLNGDLDGDFGPHGIGTSGLOIRMGQITNSFOLY
VAIGGLVLAGLEFGWPHYHRAKRAKLEGGPNEBMLNHLAVLLGCSLGMWAGLI
HVSAPNKLNDAGVSKYDIPLPHFEILNGLFLELITVGHMYRTMMGSGHSIKELNKGPT
TFKGLNPTVGTGLMLEDISHHDLIAVLFLIIGMHMYRTMMGSGHSIKELNKGPT
GGHKGLEYENMTSMHQAOLATNLAFGLSGITIVVHMYAMPYATPYATQIOLCT
HHMIGGLIYGAAGATITFMRDIDPVYNNVNDVRYRHMDATISLHMYCIPFG
HSFGLVHNDTRALGRPODMFSDTAIQDPFAQVQVHITLAPSTAPNLEPST
AGCGGIVAGGVVAMPVIALGADPLVHIIHAFTHVYLLILKGLFARNRSLIPDK
ANLGEFFPCDGRGCTCGVSCMDVFLGLFPMYNSISIVIFSKMQSDVWGTVA
ACNVTITGNNNAOSALTINGMLRDLMAQSVQVINSYGSALSAYGIMLGHIWAF
SIMPLESGSGVGGYELIESIVMHNKLVKVAIOPRALSTIQGRVAVHYLLGIATTT
NAFPAHILISLC"
2971..5202
/gene="psab"
2971..5202
/gene="psab"
/note="psab"
/note="binds chlorophylls, iron-sulfur centers and
phyllonones"
/codon_start=1
/transl_table=11
/product="photosystem I psab"
/protein_id="AAC12867.1"
/db_xref="GI:2827201"
/translation="MATKPKFSODLAODPTRTIRYAMATGNDPESHGMEENLYO
KIFATIRGVALIIFLMASSLFLRVANQFDEWIKDPHIRIRAHINDPQKRAIE
ATTOGGANPVINISYGLIYHMYTIGMRNNNDLMKAVELLILASVFLAGWHLQPK
FRPSLWPKSAERLNNHLGLAGLVSSLSMAACHLHVAIPESRGGQVGNNNLSTPH
PAGIDPFEGNAGVYASDDPTANHVSGTGGAGTALITFLGFGHGTESLMTDMMH
HLAIVIFITIGHMYRTNFGISIKEMNSAGPVCNSGOFLLPHQGYDYNN
SLHFOGLHLAGTITTSVVAOHMYAMPYATPYATQIOLATYHVOVIAIEMLGA
FAHGLFWRVDYDPEONKGVLDRLYLKKEALISHSWVSLGLFTTLLYHNDVYV
AFTGPERKQILIEPVFAQPTQASHGKVLGLVNLSPDSVATAYPNGVNLQGLAD
AINGNSLFLTIGPDLVHHAFAFALATHTVLLVKGALDARGLMPDKDQGFAD
PCDGPGRGCTCISADVATLAFNALNVGVTGTFVHMKHLGIMOGVNAQPNESSTY
LKGMPRDIYMANSAOLINGVYNNMNSVMMMLFGLHVNATGFMILSRGWQOE
LIETLVMAHERPLANLVWKKRPVALSTIVQARLVGLHTFTYGYVLTVAFLIASTAG
KFG"
BASE COUNT 1378 a 1396 c 1235 g 1559 t
ORIGIN
Query Match 92.0%; Score 18.4; DB 1; Length 5568;
Best Local Similarity 95.0%; Pred. No. 2.2e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTTGTTAGTTATTAGTT 20
|||||
Db 2924 TTAGTTGTTAGTTATTAGTT 2905

RESULT 11
AX251168 11029 bp DNA linear PAT 05-OCT-2001
LOCUS AX251168
DEFINITION Sequence 136 from Patent WO0168912.
ACCESSION AX251168
VERSION AX251168.1 GI:15984591
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with tumor suppressor genes and oncogenes
JOURNAL Patent: WO 0168912-A 136 20-SEP-2001;
Epigenomics AG (DE)
FEATURES
source
location/Qualifiers
1..11029
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 2865 a 205 c 2388 g 5571 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 11029;
Best Local Similarity 95.0%; Pred. No. 1.6e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTTGTTAGTTATTAGTT 20
|||||
Db 7283 TTAGTTGTTAGTTATTAGTT 7302

RESULT 12
AX348425 16258 bp DNA linear PAT 06-FEB-2002
LOCUS AX348425
DEFINITION Sequence 120 from Patent WO0202806.
ACCESSION AX348425
VERSION AX348425.1 GI:18614461
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Method and nucleic acids for pharmacogenomic methylation analysis
JOURNAL Patent: WO 0202806-A 120 10-JAN-2002;
Epigenomics AG (DE)
FEATURES
source
location/Qualifiers
1..16258
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 4242 a 330 c 3801 g 7885 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 16258;
Best Local Similarity 95.0%; Pred. No. 1.6e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTTGTTAGTTATTAGTT 20
|||||
Db 16029 TTAGTTGTTAGTTATTAGTT 16048

RESULT 13
AX348808

LOCUS AX348808 16258 bp DNA linear PAT 06-FEB-2002
DEFINITION Sequence 266 from Patent WO0202807.
ACCESSION AX348808
VERSION AX348808.1 GI:18614843
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
TITLE Diagnosis of diseases associated with cell signalling
JOURNAL Patent: WO 0202807-A 266 10-JAN-2002;
Epigenomics AG (DE)
FEATURES
source
location/Qualifiers
1..16258
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"

BASE COUNT 4242 a 330 c 3801 g 7885 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 5; Length 16258;
Best Local Similarity 95.0%; Pred. No. 1.6e+02;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTTGTTAGTTATTAGTT 20
|||||
Db 16029 TTAGTTGTTAGTTATTAGTT 16048

RESULT 14
AP004902/c 123078 bp DNA linear PAT 19-JUL-2002
LOCUS AP004902/c
DEFINITION Lotus japonicus genomic DNA, chromosome 2, clone: LjT04G24, complete sequence.
ACCESSION AP004902
VERSION AP004902.1 GI:21907918
KEYWORDS
SOURCE
ORGANISM
Lotus japonicus DNA, clone: LjT library clone: LjT04G24.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Loteae; Lotus.

REFERENCE
1 Kaneko,T., Nakamura,Y., Asamizu,E., Kato,T., Sato,S. and Tabata,S. Structural Analysis of a Lotus japonicus Genome. I. Sequence Features and Mapping of Sixty-six YAC clones which cover the 6.7 Mb Regions of the Genome
2 (bases 1 to 123078) Unpublished
Nakamura,Y.
DIRECT SUBMISSION
Submitted (26-MAR-2002) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research: 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynk@kazusa.or.jp, URL: http://www.kazusa.or.jp, Tel: 81-438-52-3935, Fax: 81-438-52-3934)

FEATURES
source
location/Qualifiers
1..123078
/organism="Lotus japonicus"
/db_xref="taxon:34305"
/chromosome="2"
/clone="LjT04G24"
/clone="LjT library"
/note="YAC clone: TM0060"

BASE COUNT 40721 a 22328 c 21292 g 38737 t

ORIGIN

Query Match 92.0%; Score 18.4; DB 7; Length 123078;
Best Local Similarity 95.0%; Pred. No. 81;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTATTAGTT 20
 |||||
 Db 40067 TTAGTGTAGTATTAGTT 40048

RESULT 15
 AC010232/c 165920 bp DNA linear PRI 26-FEB-2002
 LOCUS Homo sapiens chromosome 5 clone CTC-316C21, complete sequence.
 DEFINITION AC010232
 ACCESSION AC010232
 VERSION AC010232.10 GI:18921277
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 165920)
 DOE Joint Genome Institute and Stanford Human Genome Center.
 Direct Submission
 Unpublished
 2 (bases 1 to 165920)
 DOE Joint Genome Institute.
 Direct Submission
 Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 3 (bases 1 to 165920)
 DOE Joint Genome Institute and Stanford Human Genome Center.
 Direct Submission
 Submitted (26-FEB-2002) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA
 On Feb 26, 2002 this sequence version replaced gi:15290321.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 Finishing Completed at Stanford Human Genome Center
 www.shgc.stanford.edu
 Quality: Phrap Quality >=40 99.9% of Sequence:
 Estimated Total Number of Errors is 0.1.
 Location/Qualifiers
 1. 165920
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="CTC-316C21"

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

FEATURES
 source
 1. 165920
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="CTC-316C21"

BASE COUNT 45866 a 32867 c 34867 g 52500 t
 ORIGIN

Query Match 92.0%; Score 18.4; DB 8; Length 165920;
 Best Local Similarity 95.0%; Pred. No. 74;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTATTAGTT 20
 |||||
 Db 50747 TTAGTGTAGTATTAGTT 50728

Search completed: April 6, 2003, 11:57:43
 Job time : 305 secs

PA (AVER) AVENTIS PASTEUR.
XX
PI Bachy M, Sodoyer R, Tranoy E;
XX
DR WPI; 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially Interleukin-10 or Interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
CC
XX
SQ Sequence 20 BP; 4 A; 0 C; 4 G; 12 T; 0 other;
Query Match 100.0%; Score 20; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 TTAGTGTAGTTATTAGTT 20
1 TTAGTGTAGTTATTAGTT 20
Db 1 TTAGTGTAGTTATTAGTT 20
RESULT 2
AAF28872
ID AAF28872 standard; DNA: 20 BP.
XX
AC AAF28872;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #1 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW Propylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /*tag= a
FT /note= "contain phosphorothioate internucleotide
linkages"
XX
PN WO200075304-A1.
XX
PD 14-DEC-2000.
XX
PF 08-JUN-2000; 2000MO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVER) AVENTIS PASTEUR.
XX
PI Bachy M, Sodoyer R, Tranoy E;
XX
DR WPI; 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in

PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at
CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
CC oligonucleotides do not contain any CG dinucleotides in which C are
CC unethylated. The oligonucleotides are used as human immunostimulants
CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
CC They induce proliferation of human lymphocytes, induce secretion of
CC cytokines, especially Interleukin-10 or Interferon-gamma and increase
CC expression of the CD86 activation marker or the CD25 cytokine receptor
CC on human B lymphocytes.
CC
XX
SQ Sequence 20 BP; 5 A; 0 C; 3 G; 12 T; 0 other;
Query Match 92.0%; Score 18.4; DB 22; Length 20;
Best Local Similarity 95.0%; Pred. No. 97;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1 TTAGTGTAGTTATTAGTT 20
1 TTAGTGTAGTTATTAGTT 20
Db 1 TTAGTGTAGTTATTAGTT 20
RESULT 3
AAF28876
ID AAF28876 standard; DNA: 20 BP.
XX
AC AAF28876;
XX
DT 09-MAY-2001 (first entry)
XX
DE Immunostimulatory oligonucleotide #5 as vaccine adjuvant.
XX
KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
KW Propylactic vaccine; human lymphocyte proliferation; interferon-gamma;
KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
KW phosphorothioate; ss.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT modified_base 1..20
FT /*tag= a
FT /note= "contain phosphorothioate internucleotide
linkages"
XX
PN WO200075304-A1.
XX
PD 14-DEC-2000.
XX
PF 08-JUN-2000; 2000MO-FR01566.
XX
PR 08-JUN-1999; 99FR-0007457.
PR 06-AUG-1999; 99FR-0010378.
XX
PA (AVER) AVENTIS PASTEUR.
XX
PI Bachy M, Sodoyer R, Tranoy E;
XX
DR WPI; 2001-041317/05.
XX
PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
PT vaccines for human use, induce lymphocyte proliferation and cytokine
PT secretion -
XX
PS Example 4; Page 13; 30pp; French.
XX
CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
CC immunostimulatory oligonucleotide of the invention which contains at

CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
 CC oligonucleotides do not contain any CG dinucleotides in which C are
 CC unethyiated. The oligonucleotides are used as human immunostimulants
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
 CC They induce proliferation of human lymphocytes, induce secretion of
 CC cytokines, especially interleukin-10 or interferon-gamma and increase
 CC expression of the CD86 activation marker or the CD25 cytokine receptor
 CC on human B lymphocytes.

XX Sequence 20 BP: 4 A; 0 C; 3 G; 13 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;

Best Local Similarity 95.0%; Pred. No. 97;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTGTAGTTATTAGTT 20
 Db 1 TTAGTGTAGTTATTAGTT 20

RESULT 4

AAF28880 ID AAF28880 standard; DNA; 20 BP.

XX AAF28880;

DT 09-MAY-2001 (first entry)

XX Immunostimulatory oligonucleotide #9 as vaccine adjuvant.

XX Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;

KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;

KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;

XX phosphorothioate; ss.

XX Synthetic.

FT Key Location/Qualifiers
 modified_base 1..20
 FT /tag= a
 FT /note= "contain phosphorothioate internucleotide
 FT linkages"

XX WO200075304-A1.

XX 14-DEC-2000.

XX 08-JUN-2000; 2000WO-FR01566.

XX 08-JUN-1999; 99FR-0007457.

XX 06-AUG-1999; 99FR-0010376.

XX (AVET) AVENTIS PASTEUR.

XX Bachy M, Sodoyer R, Tranoy E;

XX WPI; 2001-041317/05.

XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
 PT vaccines for human use, induce lymphocyte proliferation and cytokine
 PT secretion -

XX Example 4; Page 13; 30pp; French.

CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
 CC immunostimulatory oligonucleotide of the invention which contains at
 CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
 CC oligonucleotides do not contain any CG dinucleotides in which C are
 CC unethyiated. The oligonucleotides are used as human immunostimulants
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
 CC They induce proliferation of human lymphocytes, induce secretion of
 CC cytokines, especially interleukin-10 or interferon-gamma and increase
 CC expression of the CD86 activation marker or the CD25 cytokine receptor

CC on human B lymphocytes.

XX Sequence 20 BP: 4 A; 1 C; 3 G; 12 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;

Best Local Similarity 95.0%; Pred. No. 97;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTGTAGTTATTAGTT 20
 Db 1 TTAGTGTAGTTATTAGTT 20

RESULT 5

AAF28885 ID AAF28885 standard; DNA; 20 BP.

XX AAF28885;

DT 09-MAY-2001 (first entry)

XX Immunostimulatory oligonucleotide #14 as vaccine adjuvant.

XX Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;

KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;

KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;

XX phosphorothioate; ss.

XX Synthetic.

FT Key Location/Qualifiers
 modified_base 1..20
 FT /tag= a
 FT /note= "contain phosphorothioate internucleotide
 FT linkages"

XX WO200075304-A1.

XX 14-DEC-2000.

XX 08-JUN-2000; 2000WO-FR01566.

XX 08-JUN-1999; 99FR-0007457.

XX 06-AUG-1999; 99FR-0010376.

XX (AVET) AVENTIS PASTEUR.

XX Bachy M, Sodoyer R, Tranoy E;

XX WPI; 2001-041317/05.

XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
 PT vaccines for human use, induce lymphocyte proliferation and cytokine
 PT secretion -

XX Example 4; Page 13; 30pp; French.

CC Oligonucleotides AAF28872-AAF28887 represent specific examples of an
 CC immunostimulatory oligonucleotide of the invention which contains at
 CC least one sequence 5'-TTNIN2TT-3' where N1 and N2 are A, T, C or G. The
 CC oligonucleotides do not contain any CG dinucleotides in which C are
 CC unethyiated. The oligonucleotides are used as human immunostimulants
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
 CC They induce proliferation of human lymphocytes, induce secretion of
 CC cytokines, especially interleukin-10 or interferon-gamma and increase
 CC expression of the CD86 activation marker or the CD25 cytokine receptor
 CC on human B lymphocytes.

XX Sequence 20 BP: 3 A; 0 C; 4 G; 13 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;

Best Local Similarity 95.0%; Pred. No. 97;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTGTAGTTATTAAGT 20
 |||||
 Db 1 TTAGTGTAGTTATTAAGT 20

RESULT 6

AAAF28886
 ID AAF28886 standard; DNA: 20 BP.

AC AAF28886;

DT 09-MAY-2001 (first entry)

DE Immunostimulatory oligonucleotide #15 as vaccine adjuvant.

KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
 KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
 KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
 KW phosphorothioate; ss.

OS Synthetic.

Key Location/Qualifiers
 modified_base 1..20

FT /tag= a
 FT /note= "contain phosphorothioate internucleotide linkages"

PN WO200075304-A1.

PD 14-DEC-2000.

PF 08-JUN-2000; 2000WO-FR01566.

PR 08-JUN-1999; 99FR-0007457.

PR 06-AUG-1999; 99FR-0010378.

PA (AVET) AVENTIS PASTEUR.

PI Bachy M, Sodayer R, Tranroy E;

DR WPI; 2001-041317/05.

XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
 PT vaccines for human use, induce lymphocyte proliferation and cytokine
 PT secretion -

XX Example 4; Page 13; 30pp; French.

XX Oligonucleotides AAF28872-AAF28887 represent specific examples of an
 CC immunostimulatory oligonucleotide of the invention which contains at
 CC least one sequence 5'-TTN1N2TT-3' where N1 and N2 are A, T, C or G. The
 CC oligonucleotides do not contain any CG dinucleotides in which C are
 CC unmethylated. The oligonucleotides are used as human immunostimulants
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
 CC They induce proliferation of human lymphocytes, induce secretion of
 CC cytokines, especially interleukin-10 or interferon-gamma and increase
 CC expression of the CD86 activation marker or the CD25 cytokine receptor
 CC on human B lymphocytes.

XX Sequence 20 BP; 3 A; 1 C; 4 G; 12 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;

Best Local Similarity 95.0%; Pred. No. 97;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTGTAGTTATTAAGT 20

Db 1 TTAGTGTAGTTATTAAGT 20

RESULT 7

AAAF28887
 ID AAF28887 standard; DNA: 20 BP.

AC AAF28887;

DT 09-MAY-2001 (first entry)

DE Immunostimulatory oligonucleotide #16 as vaccine adjuvant.

KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
 KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
 KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor;
 KW phosphorothioate; ss.

OS Synthetic.

Key Location/Qualifiers
 modified_base 1..20

FT /tag= a
 FT /note= "contain phosphorothioate internucleotide linkages"

PN WO200075304-A1.

PD 14-DEC-2000.

PF 08-JUN-2000; 2000WO-FR01566.

PR 08-JUN-1999; 99FR-0007457.

PR 06-AUG-1999; 99FR-0010378.

PA (AVET) AVENTIS PASTEUR.

PI Bachy M, Sodayer R, Tranroy E;

DR WPI; 2001-041317/05.

XX New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
 PT vaccines for human use, induce lymphocyte proliferation and cytokine
 PT secretion -

XX Example 4; Page 13; 30pp; French.

XX Oligonucleotides AAF28872-AAF28887 represent specific examples of an
 CC immunostimulatory oligonucleotide of the invention which contains at
 CC least one sequence 5'-TTN1N2TT-3' where N1 and N2 are A, T, C or G. The
 CC oligonucleotides do not contain any CG dinucleotides in which C are
 CC unmethylated. The oligonucleotides are used as human immunostimulants
 CC and as adjuvants in therapeutic and prophylactic vaccines for human use.
 CC They induce proliferation of human lymphocytes, induce secretion of
 CC cytokines, especially interleukin-10 or interferon-gamma and increase
 CC expression of the CD86 activation marker or the CD25 cytokine receptor
 CC on human B lymphocytes.

XX Sequence 20 BP; 3 A; 0 C; 5 G; 12 T; 0 other;

Query Match 92.0%; Score 18.4; DB 22; Length 20;

Best Local Similarity 95.0%; Pred. No. 97;

Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTGTAGTTATTAAGT 20

Db 1 TTAGTGTAGTTATTAAGT 20

RESULT 8

ABO21728
 ID ABO21728 standard; DNA: 749 BP.

AC ABO21728;

DT 12-JUL-2002 (first entry)

DE oligonucleotide for detecting cytosine methylation SEQ ID NO 8319.
 XX Human: cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;
 KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200218632-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 01-SEP-2001; 2001WO-EP10074.
 XX
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PR 05-SEP-2000; 2000DE-1044543.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX
 DR WPI; 2002-371829/40.
 XX
 PT Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 CC
 XX
 SQ Sequence 749 BP; 173 A; 54 C; 199 G; 322 T; 1 other;
 XX
 Query Match 92.0%; Score 18.4; DB 24; Length 749;
 Best Local Similarity 95.0%; Pred. NO. 80;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 TTAGTGTGTTAGTTATTAGTT 20
 ||||| ||||| ||||| |||||
 DB 347 TTAGTTTATTAGTTATTAGTT 366
 XX
 RESULT 9
 AB021729/c
 ID AB021729 standard; DNA; 749 BP.
 XX
 AC AB021729;
 XX
 DT 12-JUL-2002 (first entry)
 XX
 DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 8320.
 XX
 KW Human: cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;

KW drug; side effect; cancer; central nervous system; cardiovascular;
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;
 KW SNP; cell differentiation; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO200218632-A2.
 XX
 PD 07-MAR-2002.
 XX
 PF 01-SEP-2001; 2001WO-EP10074.
 XX
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PR 05-SEP-2000; 2000DE-1044543.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K, Guetig D;
 XX
 DR WPI; 2002-371829/40.
 XX
 PT Determining the degree of cytosine methylation in genomic DNA, useful
 PT for diagnosis and prognosis, comprises selective hybridization of
 PT amplicons from chemically treated DNA -
 XX
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.
 XX
 XX This invention describes a novel method for determining the degree of
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
 CC genomic sample of DNA. The sample is treated chemically to convert
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic
 CC DNA that contains the target C is amplified to form a labeled amplicon.
 CC The amplicon is hybridised to two classes, each with at least one
 CC member, of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers
 CC and the degree of hybridisation to both classes is determined from the
 CC label on the amplicon. From the ratio of labels hybridised to the two
 CC classes of oligomers, the degree of methylation is calculated. The method
 CC is used: (i) for diagnosis and/or prognosis of side effects of
 CC therapeutic drugs and of a wide range of diseases, e.g. cancer, disorders
 CC of the central nervous, cardiovascular, gastrointestinal and respiratory
 CC systems etc., particularly by detecting mutations or single nucleotide
 CC polymorphisms (SNP's); and (ii) for differentiation of cell or tissue
 CC types and for investigating cell differentiation. The method allows the
 CC methylation status of many C residues to be determined simultaneously.
 CC AB013410-AB054121 represent genomic DNA sequences used to illustrate the
 CC method for determining the degree of cytosine methylation described in
 CC the disclosure of the invention.
 CC
 XX
 SQ Sequence 749 BP; 322 A; 199 C; 54 G; 173 T; 1 other;
 XX
 Query Match 92.0%; Score 18.4; DB 24; Length 749;
 Best Local Similarity 95.0%; Pred. NO. 80;
 Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 OY 1 TTAGTGTGTTAGTTATTAGTT 20
 ||||| ||||| ||||| |||||
 DB 403 TTAGTTTATTAGTTATTAGTT 384
 XX
 RESULT 10
 AAS46414
 ID AAS46414 standard; DNA; 11029 BP.
 XX
 AC AAS46414;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Tumour suppressor gene derived chemically modified sequence #136.
 XX
 KW Human: tumour suppressor gene; oncogene; antitumour; cytostatic;
 KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
 KW cytosine methylation; ds.
 XX

```
OS Homo sapiens.
XX
PN WO200168912-A2.
XX
PD 20-SEP-2001.
XX
PF 15-MAR-2001; 2001WO-EP02955.
XX
PR 15-MAR-2000; 2000DE-1013847.
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2001-602752/68.
XX
PT Fragments of chemically modified genes associated with tumour suppressor
PT genes and oncogenes, useful in designing primers and probes for
PT analysing diseases associated with cytosine methylation state e.g.
PT cancer
XX
PS Claim 1; SEQ ID NO 136; 27pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of 18
CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
CC bisulphite, of genes associated with tumour suppression and
CC oncogenes having a sequence taken from 536 (actually 533 since
CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
CC (5s) and sequences complementary to (5s). The nucleic acid may be a
CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
CC form part of a set of probes for detecting the cytosine methylation state
CC and/or single nucleotide polymorphisms and also to be used in an
CC array for analysing diseases associated with CpG dinucleotides e.g.
CC cancers and tumours. The probes can also be used in a method for
CC ascertaining genetic and/or epigenetic parameters for the diagnosis
CC and/or therapy of existing diseases or the predisposition to specific
CC diseases, by analysing cytosine methylations. The parameters may be
CC compared to another set of genetic and/or epigenetic parameters, the
CC differences serving as basis for diagnosis and/or prognosis events which
CC are disadvantageous to patients. The present sequence is one of the
CC 533 genomic sequences derived from tumour suppressor genes and
CC oncogenes. Sequences with even numbered Seq ID numbers are the
CC complementary sequence of the corresponding odd numbered sequence (e.g.
CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
CC is missing).
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 11029 BP; 2865 A; 205 C; 2388 G; 5571 T; 0 other;

Query Match          92.0%; Score 18.4; DB 22; Length 11029;
Best Local Similarity 95.0%; Pred. No. 70;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTGTAGTTATTAGTT 20
   ||||||||||||||||
DB 7283 TTAGTGTAGTTATTATTT 7302

RESULT 11
ABL70376
ID ABL70376 standard; DNA; 16258 BP.
XX
AC ABL70376;
XX
DF 01-JUL-2002 (first entry)
XX
```

```
DE Chemically treated cell signalling DNA sequence complementary to#133.
XX
XX Cell signalling; cytosine methylation; cell signalling disease;
XX cancer; tumour; cytostatic; ds.
XX
XX Unidentified.
XX
PN WO200202807-A2.
XX
PD 10-JAN-2002.
XX
PF 29-JUN-2001; 2001WO-EP07471.
XX
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-154758/20.
XX
PT Nucleic acid, useful for diagnosis and therapy of diseases associated
PT with cell signalling e.g. cancer, comprises chemically modified genomic
PT sequences of genes associated with cell signalling
XX
PS Claim 1; SEQ ID NO 266; 24pp+sequence listing; English.
XX
CC The invention relates to a nucleic acid comprising a sequence of at least
CC 18 bases of a segment of chemically pretreated DNA of genes associated
CC with cell signalling. The activity of the modified sequences of the
CC invention may be described as cytostatic. The object of the invention is
CC to provide the chemically modified DNA of genes associated with cell
CC signalling, as well as oligonucleotides and/or PNA-oligomers for
CC detecting cytosine methylations, as well as a method which is
CC particularly suitable for the diagnosis and/or therapy of genetic and
CC epigenetic parameters of genes associated with cell signalling. The
CC chemically modified DNA provided by the invention is useful for diagnosis
CC and therapy of diseases such as solid tumours and cancer. The sequences
CC given in records ABL70111-ABL70626 represent chemically pre-treated
CC genomic DNA's of genes associated with cell signalling.
CC Note: The sequence data for this patent is not represented in the printed
CC specification, but is based on sequence information supplied by the
CC European Patent Office.
XX
SQ Sequence 16258 BP; 4242 A; 330 C; 3801 G; 7885 T; 0 other;

Query Match          92.0%; Score 18.4; DB 24; Length 16258;
Best Local Similarity 95.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 TTAGTGTAGTTATTAGTT 20
   ||||||||||||||||
DB 16029 TTAGTGTAGTTTATTAGTT 16048

RESULT 12
ABK40038
ID ABK40038 standard; DNA; 16258 BP.
XX
AC ABK40038;
XX
DT 21-MAY-2002 (first entry)
XX
DE Human chemically pretreated gene sequence #60 strand 2.
XX
XX Human; ds; bisulphite treatment; CPG; DNA methylation; cancer; tumour;
XX cytostatic; ALDH6; CYP11A; CYP11B; CYP3A3; DPTD; EPHX2; OCLN; TXNRD1;
XX UGFB; MRP; pharmacogenomics; SNP; single nucleotide polymorphism.
XX
OS Homo sapiens.
XX
PN WO200202806-A2.
```

XX 10-JAN-2002.
PD
XX
PF 29-JUN-2001; 2001WO-EP07470.
XX
PR 30-JUN-2000; 2000DE-1032529.
XX
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPiG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
DR WPI; 2002-154757/20.
XX
PI New nucleic acid, oligonucleotides and peptide nucleic acid-oligomers,
PI useful for detecting cytosine methylation state of genes associated
PI with pharmacogenomics and for therapy of diseases e.g. cancer
XX
PS Claim 1; SEQ ID No 120; 24pp; English.
XX
CC The invention relates to a nucleic acid comprising a sequence at
CC least 18 bases in length of a segment of the chemically pretreated DNA
CC of genes associated with pharmacogenomics according to one of the
CC sequences of the genes ALDH6 (NM_000693), CYP11A (NM_000781), CYP11B1
CC (NM_000497), CYP3A3 (NM_000776 and NM_017460), DYPD (NM_000110), EPHX2
CC (NM_001979), OCLN (NM_002538), TXNRD1 (NM_003330), UGT8 (NM_003360),
CC MRP (NM_004996, NM_019900, NM_019901, NM_019902, NM_019862, NM_019898,
CC NM_019899) and their complementary sequences, or a sequence (SI) chosen
CC from 87 sequences and their complementary sequences. The chemical pretreatment
CC is disulphite treatment to convert cytosines (but not methyl-cytosines)
CC into uracils. Also included are an oligomer (II) in particular an
CC oligonucleotide or a peptide nucleic acid (PNA)-oligomer, comprising in
CC each case at least one base sequence having a length of 9 nucleotides
CC which hybridises to or is identical to a chemically pretreated DNA of
CC genes associated with pharmacogenomics and their complements; arranged in
CC an array for analysing diseases associated with the polymorphisms)
CC (CpG) and/or detecting SNPs (single nucleotide polymorphisms)
CC of the 87 sequences. The oligomers may also be used as PCR primers.
CC The set of 87 nucleic acids and their complements is useful for diagnosis
CC and therapy of solid tumours and cancer. The present sequence
CC represents one the 87 DNA sequences or its complement.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 16258 BP; 4242 A; 330 C; 3801 G; 7885 T; 0 other;
XX
Query Match 92.0%; Score 18.4; DB 24; Length 16258;
Best Local Similarity 95.0%; Pred. No. 68;
Matches 19; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 TTAGTTGTTAGTTATTAGTT 20
IIIIIIIIIIIIIIIIIIII
Db 16029 TTAGTTGTTAGTTTATTAGTT 16048
XX
RESULT 13
AAK65476/C
ID AAK65476 standard; DNA; 27435 BP.
XX
AC AAK65476;
XX
XX 06-NOV-2001 (first entry)
XX
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:20288.
XX
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX
OS Homo sapiens.
XX
XX WO200157182-A2.
PN

XX 09-AUG-2001.
PD
XX
PF 17-JAN-2001; 2001WO-US01354.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205155.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226682.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228824.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR

PR 29-SEP-2000; 2000US-0236367.
 PR 29-SEP-2000; 2000US-0236368.
 PR 29-SEP-2000; 2000US-0236369.
 PR 29-SEP-2000; 2000US-0236370.
 PR 02-OCT-2000; 2000US-0236802.
 PR 02-OCT-2000; 2000US-0237037.
 PR 02-OCT-2000; 2000US-0237038.
 PR 02-OCT-2000; 2000US-0237039.
 PR 13-OCT-2000; 2000US-0239935.
 PR 13-OCT-2000; 2000US-0239937.
 PR 20-OCT-2000; 2000US-0240960.
 PR 20-OCT-2000; 2000US-0241221.
 PR 20-OCT-2000; 2000US-0241785.
 PR 20-OCT-2000; 2000US-0241786.
 PR 20-OCT-2000; 2000US-0241787.
 PR 20-OCT-2000; 2000US-0241808.
 PR 20-OCT-2000; 2000US-0241809.
 PR 01-NOV-2000; 2000US-0244617.
 PR 08-NOV-2000; 2000US-0246474.
 PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249297.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250381.
 PR 01-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Rosen CA, Barash SC, Ruben SM;
 XX

DR WPI; 2001-483426/52.
 XX
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
 PT useful for preventing, diagnosing and/or treating cancers and
 PT metastasis -
 XX
 PS Disclosure; SEQ ID NO 20288; 3071pp + Sequence Listing; English.
 XX
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)
 CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
 CC activity, and can be used in gene therapy and vaccine production. (I)
 CC proteins and polynucleotides may be used in the prevention, diagnosis and
 CC treatment of diseases associated with inappropriate (I) expression. For
 CC example, they may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of (I) by expressing inactive proteins or to
 CC supplement the patients own production of (I). Additionally, (I)
 CC polynucleotides may be used to produce the secreted (I), by inserting
 CC the nucleic acids into a host cell and culturing the cell to express the
 CC protein. (I) proteins and polynucleotides may be used to prevent,
 CC diagnose and treat immune/hematopoietic-related diseases, especially
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703
 CC to AAK87694 represent human immune/hematopoietic antigen genomic
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169
 CC represent sequences used in the exemplification of the present invention.
 XX
 SQ Sequence 27435 BP; 7241 A; 5321 C; 5251 G; 9622 T; 0 other:
 XX
 Query Match 90.0%; Score 18; DB 22; Length 27435;
 Best Local Similarity 100.0%; Pred. No. 96;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 TTAGTGTAGTATATAG 18
 Db 1648 TTACTGTAGTATATAG 1631
 ||||||||||||||||
 ID AAF28871 standard; DNA; 20 BP.
 XX AAF28871:
 AC AAF28871:
 XX
 DT 09-MAY-2001 (first entry)
 XX
 DE Generic immunostimulatory oligonucleotide as vaccine adjuvant.
 XX
 KW Immunostimulant; immunostimulatory; human; adjuvant; interleukin-10;
 KW prophylactic vaccine; human lymphocyte proliferation; interferon-gamma;
 KW cytokine secretion inducer; CD86 activation; CD25 cytokine receptor; ss.
 XX
 OS Synthetic.
 XX
 PN WO200075304-A1.
 PD 14-DEC-2000.
 XX
 PF 08-JUN-2000; 2000WO-EP01566.
 XX
 PR 08-JUN-1999; 99FR-0007457.
 PR 06-AUG-1999; 99FR-0010378.
 XX
 PA (AVER) AVENTIS PASTEUR.
 XX
 PI Bachy M, Sodayer R, Tranney E;
 XX
 DR WPI; 2001-041317/05.
 XX
 PT New immunostimulatory oligonucleotides, useful e.g. as adjuvants in
 PT vaccines for human use, induce lymphocyte proliferation and cytokine
 PT secretion -
 XX
 PS Claim 8; Page 17; 30pp; French.

XX This sequence represents a generic example of an immunostimulatory
 CC oligonucleotide of the invention which contains at least one sequence
 CC 5'-TTTNN2TT-3' where N1 and N2 are A, T, C or G. The oligonucleotides
 CC do not contain any CG dinucleotides in which C are unmethylated.
 CC The oligonucleotides are used as human immunostimulants and as adjuvants
 CC in therapeutic and prophylactic vaccines for human use. They induce
 CC proliferation of human lymphocytes, induce secretion of cytokines,
 CC especially interleukin-10 or interferon-gamma and increase expression
 CC of the CD86 activation marker or the CD25 cytokine receptor on human B
 CC lymphocytes.
 XX
 SQ Sequence 20 BP; 3 A; 1 C; 3 G; 12 T; 1 other;

Query Match 87.0%; Score 17.4; DB 22; Length 20;
 Best Local Similarity 90.0%; Pred. No. 2.5e+02;
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTTGTTAGTTATTAGTT 20
 ||||| ||||| |||||
 Db 1 TTAGTTCTTAGTTTNTTACTT 20

RESULT 15
 ABL13576/c
 ID ABL13576 standard; cDNA: 2526 BP.
 XX
 AC ABL13576;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 35210.
 XX
 KW Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical; gene; ss.
 XX
 OS Drosophila melanogaster.
 XX
 PN W0200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI: 2001-656860/75.
 DR P-PSDB; ABB69473.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Claim 1; SEQ ID NO 35210; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins
 CC sequences (ABB57737-ABB72072).
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 2526 BP; 759 A; 488 C; 513 G; 766 T; 0 other;

Query Match 87.0%; Score 17.4; DB 23; Length 2526;
 Best Local Similarity 94.7%; Pred. No. 1.9e+02;
 Matches 18; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 TTAGTTGTTAGTTATTAGTT 19
 ||| ||||| ||||| |||||
 Db 843 TTAATTGTTAGTTATTAGTT 825

Search completed: April 6, 2003, 12:06:03
 Job Time : 163 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:12:37 ; Search time 45.25 Seconds
(without alignments)
135.548 Million cell updates/sec

Title: US-09-980-265-12

Perfect score: 20
Sequence: 1 ttagtctgattgattgattg 20

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued_patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCrUS_COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16.4	82.0	348	1	US-08-220-606B-54 Sequence 54, Appl
2	15.8	79.0	5526	3	US-08-751-359-21 Sequence 21, Appl
3	15.8	79.0	5526	4	US-08-907-146-21 Sequence 21, Appl
4	15.8	79.0	17327	1	US-07-906-871-15 Sequence 15, Appl
5	15.2	76.0	1578	4	US-09-134-001C-677 Sequence 677, App
6	15.2	76.0	2509	1	US-08-047-413-14 Sequence 14, Appl
7	15.2	76.0	2509	3	US-08-229-050-14 Sequence 14, Appl
8	15.2	76.0	2509	3	US-08-801-563-14 Sequence 14, Appl
9	15.2	76.0	2823	1	US-08-398-008A-1 Sequence 1, Appl
10	15.2	76.0	2823	2	US-08-893-333-1 Sequence 1, Appl
11	15.2	76.0	2862	1	US-08-148-209A-1 Sequence 1, Appl
12	15.2	76.0	4483	1	US-08-181-271A-5 Sequence 5, Appl
13	15.2	76.0	4483	1	US-08-449-315-5 Sequence 5, Appl
14	15.2	76.0	4483	1	US-08-444-803-5 Sequence 5, Appl
15	15.2	76.0	4483	1	US-08-449-043-5 Sequence 5, Appl
16	15.2	76.0	4483	1	US-08-456-262A-5 Sequence 5, Appl
17	15.2	76.0	4483	1	US-08-456-265A-5 Sequence 5, Appl
18	15.2	76.0	4483	1	US-08-455-244-5 Sequence 5, Appl
19	15.2	76.0	4483	1	US-08-454-876-5 Sequence 5, Appl
20	15.2	76.0	4483	2	US-08-457-364-5 Sequence 5, Appl
21	15.2	76.0	4483	2	US-08-456-262-5 Sequence 5, Appl
22	15.2	76.0	4483	2	US-08-456-240-5 Sequence 5, Appl
23	15.2	76.0	4483	2	US-08-455-736-5 Sequence 5, Appl
24	15.2	76.0	4483	2	US-08-971-217-5 Sequence 5, Appl
25	15.2	76.0	4483	4	US-09-350-600-5 Sequence 5, Appl
26	15.2	76.0	4699	1	US-08-181-271A-6 Sequence 6, Appl
27	15.2	76.0	4699	1	US-08-449-315-6 Sequence 6, Appl

28	15.2	76.0	4699	1	US-08-444-803-6 Sequence 6, Appl
29	15.2	76.0	4699	1	US-08-449-043-6 Sequence 6, Appl
30	15.2	76.0	4699	1	US-08-456-265A-6 Sequence 6, Appl
31	15.2	76.0	4699	1	US-08-455-416-6 Sequence 6, Appl
32	15.2	76.0	4699	1	US-08-453-244-6 Sequence 6, Appl
33	15.2	76.0	4699	1	US-08-454-876-6 Sequence 6, Appl
34	15.2	76.0	4699	2	US-08-457-364-6 Sequence 6, Appl
35	15.2	76.0	4699	2	US-08-456-262-6 Sequence 6, Appl
36	15.2	76.0	4699	2	US-08-456-240-6 Sequence 6, Appl
37	15.2	76.0	4699	2	US-08-455-736-6 Sequence 6, Appl
38	15.2	76.0	4699	2	US-08-971-217-6 Sequence 6, Appl
39	15.2	76.0	4699	4	US-09-350-600-6 Sequence 6, Appl
40	15.2	76.0	59065	4	US-09-813-817-3 Sequence 3, Appl
41	15.2	76.0	59065	4	US-09-978-197-3 Sequence 3, Appl
42	14.8	74.0	200	2	US-08-875-972-27 Sequence 27, Appl
43	14.8	74.0	203	4	US-09-605-785-717 Sequence 717, App
44	14.8	74.0	292	1	US-08-644-664B-41 Sequence 41, Appl
45	14.8	74.0	292	2	US-08-761-277A-41 Sequence 41, Appl

ALIGNMENTS

RESULT 1
US-08-220-606B-54
Sequence 54, Application US/08220606B
Patent No. 5641661
GENERAL INFORMATION:
APPLICANT: Kumagai, Monto H.
APPLICANT: Genadi, Sverlow J.
TITLE OF INVENTION: P1chia Pastoris Alcohol Oxidase Z2A1 and
TITLE OF INVENTION: Z2A2 Regulatory Regions for Heterologous Gene Expression
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/220,606B
FILING DATE: 25-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,227
REFERENCE/DOCKET NUMBER: 8129-065
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-854-3660
TELEFAX: 415-854-3694
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 54:
SEQUENCE CHARACTERISTICS:
LENGTH: 348 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: DNA
US-08-220-606B-54
Query Match 82.0%; Score 16.4; DB 1; Length 348;
Best Local Similarity 94.4%; Pred. No. 54;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 2 TAGTGTAGTATTAGT 19
Db 168 TAGTGTAGTATTAGT 185

RESULT 2
US-08-751-359-21
Sequence 21, Application US/08751359
Patent No. 6143559
GENERAL INFORMATION:
APPLICANT: Michael, Nancy M
APPLICANT: Accavitti, Marianne
APPLICANT: Thompson, Craig B
TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF CHICKEN
TITLE OF INVENTION: MONOCLONAL ANTIBODIES
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESS: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/751,359
FILING DATE: Concurrently Herewith
CLASSIFICATION: 510
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: ARSB:504
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 5526 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-751-359-21
Query Match 79.0%; Score 15.8; DB 3; Length 5526;
Best Local Similarity 89.5%; Pred. No. 92;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TTAGTGTGTAGTTAGT 19
DB 3120 TTAGTGTGTAGTTAGT 3138
RESULT 3
US-08-907-146-21
Sequence 21, Application US/08907146
Patent No. 6316600
GENERAL INFORMATION:
APPLICANT: Michael, Nancy M
APPLICANT: Accavitti, Marianne
APPLICANT: Thompson, Craig B
TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF CHICKEN
TITLE OF INVENTION: MONOCLONAL ANTIBODIES
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESS: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: Texas
COUNTRY: USA
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/907,146
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/751,359
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Highlander, Steven L
REGISTRATION NUMBER: 37,642
REFERENCE/DOCKET NUMBER: ARSB:504
TELECOMMUNICATION INFORMATION:
TELEPHONE: 512/418-3000
TELEFAX: 512/474-7577
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 5526 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-907-146-21
Query Match 79.0%; Score 15.8; DB 4; Length 5526;
Best Local Similarity 89.5%; Pred. No. 92;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 TTAGTGTGTAGTTAGT 19
DB 3120 TTAGTGTGTAGTTAGT 3138
RESULT 4
US-07-906-871-15
Sequence 15, Application US/07906871
Patent No. 5340739
GENERAL INFORMATION:
APPLICANT: Stevens, Richard L.
APPLICANT: Abraham, Shalom
TITLE OF INVENTION: HEMATOPOIETIC CELL SPECIFIC
TITLE OF INVENTION: TRANSCRIPTIONAL REGULATORY ELEMENTS OF SRRGLXCIN AND USES
TITLE OF INVENTION: THEREOF
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESS: Sterne, Kessler, Goldstein & Fox
STREET: 1225 Connecticut Avenue, N.W., Suite 300
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/906,871
FILING DATE: 19920103
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/816,289
FILING DATE: 03 JAN 1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/635,544
FILING DATE: 18-JAN-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US89/03051
FILING DATE: 13-JUL-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/224,035
FILING DATE: 13-JUL-1988

ATTORNEY/AGENT INFORMATION:
NAME: Cimbalà, Michele A
REGISTRATION NUMBER: 33,851
REFERENCE/DOCKET NUMBER: 0627.2830004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)833-7533
TELEFAX: (202)833-8716
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 17327 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: both
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: exon
LOCATION: 621..753
FEATURE:
NAME/KEY: intron
LOCATION: 754..9596
FEATURE:
NAME/KEY: exon
LOCATION: 9597..9744
FEATURE:
NAME/KEY: intron
LOCATION: 9745..16396
FEATURE:
NAME/KEY: exon
LOCATION: 16397..17327
US-07-906-871-15

Query Match 79.0%; Score 15.8; DB 1; Length 17327;
Best Local Similarity 89.5%; Pred. No. 90;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTGTGTTAGTTATTAGT 19
DB 15437 TTAGTGTGTTAGTTGTTGT 15455

RESULT 5
US-09-134-001C-677
Sequence 677, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134.001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 677
LENGTH: 1578
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-677

Query Match 76.0%; Score 15.2; DB 4; Length 1578;
Best Local Similarity 85.0%; Pred. No. 1.7e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTGTGTTAGTTATTAGT 20
DB 796 TTGTATTAGTTATTGTT 815

RESULT 6
US-08-047-413-14

Sequence 14, Application US/08047413
Patent No. 5670706
GENERAL INFORMATION:
APPLICANT: Cornelissen, Bernardus J.C.
APPLICANT: Melchers, Leo S.
APPLICANT: Meulenhoff, Elisabeth J.S.
APPLICANT: van Roekel, Jeroen S.C.
APPLICANT: Sela-Buurlage, Marianne B.
APPLICANT: Vloemans, Alexandra A.
APPLICANT: Moloshuk, Charles P.
APPLICANT: Bol, John F.
TITLE OF INVENTION: FUNGAL RESISTANT PLANTS, PROCESS FOR
TITLE OF INVENTION: OBTAINING FUNGAL RESISTANT PLANTS AND RECOMBINANT
TITLE OF INVENTION: POLYNUCLEOTIDES FOR USE THEREIN
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: Morrison & Foerster
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/047,413
FILING DATE: 19-APR-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/647,831
FILING DATE: 29-JAN-1991
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H.
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 24615-20022.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-813-5600
TELEFAX: 415-494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2509 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
FEATURE:
NAME/KEY: CDS
LOCATION: join(85..142, 930..1948)
US-08-047-413-14

Query Match 76.0%; Score 15.2; DB 1; Length 2509;
Best Local Similarity 85.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTGTGTTAGTTATTAGT 20
DB 215 TTAGTGTGTTAGTTATTGCT 234

RESULT 7
US-08-229-050-14
Sequence 14, Application US/08229050
Patent No. 6066491
GENERAL INFORMATION:
APPLICANT: Cornelissen, Bernardus J.C.
APPLICANT: Melchers, Leo S.
APPLICANT: Meulenhoff, Elisabeth J.S.
APPLICANT: van Roekel, Jeroen S.C.
APPLICANT: Sela-Buurlage, Marianne B.

```

? APPLICANT: Vloemans, Alexandra A.
? APPLICANT: Woloshuk, Charles P.
? APPLICANT: Bol, John F.
? APPLICANT: Linthorst, Hubertus J.M.
? TITLE OF INVENTION: FUNGAL RESISTANT PLANTS, PROCESS FOR
? TITLE OF INVENTION: OBTAINING FUNGAL RESISTANT PLANTS AND RECOMBINANT
? NUMBER OF SEQUENCES: 17
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Morrison & Foerster
? STREET: 755 Page Mill Road
? CITY: Palo Alto
? STATE: California
? COUNTRY: USA
? ZIP: 94304-1018
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentln Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/229,050
? FILING DATE:
? CLASSIFICATION: 800
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/647,831
? FILING DATE: 29-JAN-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Murashige, Kate H.
? REGISTRATION NUMBER: 29,959
? REFERENCE/DOCKET NUMBER: 24615-20022.00
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-813-5600
? TELEFAX: 415-494-0792
? TELEX: 706141
? INFORMATION FOR SEQ. ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2509 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? FEATURE:
? NAME/KEY: CDS
? LOCATION: join(85..142, 930..1948)
?
US-08-229-050-14
?
Query Match 76.0%; Score 15.2; DB 3; Length 2509;
Best Local Similarity 85.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTGTAGTATTAGTT 20
Db 215 TTAGTGTAGTATTAGTTGCT 234

RESULT 8
US-08-801-563-14
? Sequence 14, Application US/08801563
? Patent No. 6087560
? GENERAL INFORMATION:
? APPLICANT: Cornelissen, Bernardus J.C.
? APPLICANT: Meijchers, Leo S.
? APPLICANT: Meulenhoff, Elisabeth J.S.
? APPLICANT: van Roekel, Jeroen S.C.
? APPLICANT: Sela-Buurlage, Marianne B.
? APPLICANT: Vloemans, Alexandra A.
? APPLICANT: Woloshuk, Charles P.
? APPLICANT: Bol, John F.
? APPLICANT: Linthorst, Hubertus J.M.
? TITLE OF INVENTION: FUNGAL RESISTANT PLANTS, PROCESS FOR
? TITLE OF INVENTION: OBTAINING FUNGAL RESISTANT PLANTS AND RECOMBINANT
? NUMBER OF SEQUENCES: 17
```

```

? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Morrison & Foerster
? STREET: 755 Page Mill Road
? CITY: Palo Alto
? STATE: California
? COUNTRY: USA
? ZIP: 94304-1018
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentln Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/801,563
? FILING DATE:
? CLASSIFICATION: 800
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 08/047,413
? FILING DATE: 19-APR-1993
? APPLICATION NUMBER: US 07/647,831
? FILING DATE: 29-JAN-1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Murashige, Kate H.
? REGISTRATION NUMBER: 29,959
? REFERENCE/DOCKET NUMBER: 24615-20022.00
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 415-813-5600
? TELEFAX: 415-494-0792
? TELEX: 706141
? INFORMATION FOR SEQ. ID NO: 14:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2509 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? FEATURE:
? NAME/KEY: CDS
? LOCATION: join(85..142, 930..1948)
?
US-08-801-563-14
?
Query Match 76.0%; Score 15.2; DB 3; Length 2509;
Best Local Similarity 85.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTGTAGTATTAGTT 20
Db 215 TTAGTGTAGTATTAGTTGCT 234

RESULT 9
US-08-398-008A-1
? Sequence 1, Application US/08398008A
? Patent No. 5665588
? GENERAL INFORMATION:
? APPLICANT: Kornbluth, Jack I
? TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated
? NUMBER OF SEQUENCES: 17
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Gilbreth & Adler, P.C.
? STREET: 8011 Candle Lane
? CITY: Houston
? STATE: Texas
? COUNTRY: USA
? ZIP: 77071
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette, 3.5 inch, 1.44 Mb storage
? OPERATING SYSTEM: Macintosh
? SOFTWARE: Microsoft Word 5.1
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/398,008A
? FILING DATE: March 2, 1995
```

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/126,501
FILING DATE: 24-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Dr. Benjamin Aaron
REGISTRATION NUMBER: 35,423
REFERENCE/DOCKET NUMBER: D5705CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (713) 777-2321
TELEFAX: (713) 777-6908
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2823 base pairs
TYPE: nucleic acid
STRANDEDNESS: double-stranded
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
HYPOTHETICAL: no
ANTI-SENSE: no
US-08-398-008A-1

Query Match 76.0%; Score 15.2; DB 1; Length 2823;
Best Local Similarity 85.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTACTT 20
||||| | | | | | | | | |
Db 2654 TTAGTTTAAATTATTACTT 2673

RESULT 10
US-08-893-333-1
Sequence 1, Application US/08893333A
Patent No. 5981705
GENERAL INFORMATION:
APPLICANT: Kornbluth, Jacki
TITLE OF INVENTION: DNA Encoding Natural Killer Lytic Associated Protein
FILE REFERENCE: D5705CIP/D
CURRENT APPLICATION NUMBER: US/08/893.333A
CURRENT FILING DATE: 1997-07-16
NUMBER OF SEQ ID NOS: 17
SEQ ID NO 1
LENGTH: 2823
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
LOCATION: 190..1953
OTHER INFORMATION: CDS
US-08-893-333-1

Query Match 76.0%; Score 15.2; DB 2; Length 2823;
Best Local Similarity 85.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTACTT 20
||||| | | | | | | | | |
Db 2654 TTAGTTTAAATTATTACTT 2673

RESULT 11
US-08-148-209A-1/c
Sequence 1, Application US/08148209A
Patent No. 5556780
GENERAL INFORMATION:
APPLICANT: Dzeu, Victor J
APPLICANT: Nakoyama, Masashi
TITLE OF INVENTION: TYPE-2 ANGIOTENSIN II RECEPTOR AND GENE
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Emdarcadero Center, Suite 3400

CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/148,209A
FILING DATE: 05-NOV-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A-58491-1/BIR
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2862-base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 132..1223
US-08-148-209A-1

Query Match 76.0%; Score 15.2; DB 1; Length 2862;
Best Local Similarity 85.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTACTT 20
||||| | | | | | | | | |
Db 1507 TTGTGTTAAATTATAGTT 1488

RESULT 12
US-08-181-271A-5
Sequence 5, Application US/08181271A
Patent No. 5614395
GENERAL INFORMATION:
APPLICANT: Ryals, John A.
APPLICANT: Alexander, Danny C.
APPLICANT: Beck, James J.
APPLICANT: Duesing, John H.
APPLICANT: Friedrich, Leslie B.
APPLICANT: Goodman, Robert M.
APPLICANT: Harms, Christian
APPLICANT: Meins, Jr., Frederick
APPLICANT: Montoya, Alice
APPLICANT: Moyer, Mary B.
APPLICANT: Neuhaus, Jean-Marc
APPLICANT: Payne, George B.
APPLICANT: Sperison, Christoph
APPLICANT: Stinson, Jeffrey R.
APPLICANT: Uknes, Scott J.
APPLICANT: Ward, Eric R.
APPLICANT: Williams, Shericca C.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA

ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/181,271A
FILING DATE: 13-JAN-94
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506
FILING DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672
FILING DATE: 20-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: S-19825/PL/GC 1727
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4483 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-181-271A-5

Query Match 76.0%; Score 15.2; DB 1; Length 4483;
Best Local Similarity 85.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TTAGTTGTAGTTATTGCTT 20

DB 1627 TTAGTTGATAGTTATTGCTT 1646
||||| |||||||
RESULT 13
US-08-449-315-5
Sequence 5, Application US/08449315
Patent No. 5650505
GENERAL INFORMATION:
APPLICANT: Ryals, John A.
APPLICANT: Alexander, Danny C.
APPLICANT: Beck, James J.
APPLICANT: Duesting, John H.
APPLICANT: Friedrich, Leslie B.
APPLICANT: Goodman, Robert M.
APPLICANT: Harms, Christian
APPLICANT: Meins, Jr., Frederick
APPLICANT: Montoya, Alice
APPLICANT: Moyer, Mary B.
APPLICANT: Neuhaus, Jean-Marc
APPLICANT: Payne, George B.
APPLICANT: Speisohn, Christoph
APPLICANT: Stinson, Jeffrey R.
APPLICANT: Uknes, Scott J.
APPLICANT: Ward, Eric R.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,315
FILING DATE: 24-MAY-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/181,271
FILING DATE: 13-JAN-94
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506

FILED DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILED DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILED DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672
FILED DATE: 20-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILED DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILED DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4483 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-449-315-5

Query Match 76.0%; Score 15.2; DB 1; Length 4483;
Best Local Similarity 85.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 TTAGTTGTAGTTATTAGTT 20
Db 1627 TTAGTTGATGATTATTGCT 1646

RESULT 14
US-08-444-803-5
Sequence 5, Application US/08444803
Patent No. 5654414
GENERAL INFORMATION:
APPLICANT: Ryals, John A.
APPLICANT: Alexander, Danny C.
APPLICANT: Beck, James J.
APPLICANT: Duesing, John H.
APPLICANT: Friedrich, Leslie B.
APPLICANT: Goodman, Robert M.
APPLICANT: Harms, Christian
APPLICANT: Weins, Jr., Frederick
APPLICANT: Montoya, Alice
APPLICANT: Moyer, Mary B.
APPLICANT: Neuhaus, Jean-Marc
APPLICANT: Payne, George B.
APPLICANT: Speirson, Christoph
APPLICANT: Stinson, Jeffrey R.
APPLICANT: Ukenes, Scott J.
APPLICANT: Ward, Eric R.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
NUMBER OF INVENTION: DNA SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,803
FILED DATE: 19-MAY-1995
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/181,271
FILED DATE: 13-JAN-94
APPLICATION NUMBER: US 08/093,301
FILED DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILED DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILED DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILED DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILED DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILED DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILED DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILED DATE: 20-OCT 1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/848,506
FILED DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILED DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILED DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672
FILED DATE: 20-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILED DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: S-19825/P1/CGC 1727
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4483 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-444-803-5

Query Match 76.0%; Score 15.2; DB 1; Length 4483;
Best Local Similarity 85.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTATTAGT 20
||||| |||||
Db 1627 TTAGTGTAGTATTAGT 1646

RESULT 15
US-08-449-043-5
Sequence 5, Application US/08449043
Patent No. 5689044
GENERAL INFORMATION:
APPLICANT: Ryals, John A.
APPLICANT: Alexander, Danny C.
APPLICANT: Beck, James J.
APPLICANT: Duesing, John H.
APPLICANT: Friedrich, Leslie B.
APPLICANT: Goodman, Robert M.
APPLICANT: Harms, Christian
APPLICANT: Meins, Jr., Frederick
APPLICANT: Montoya, Alice
APPLICANT: Moyer, Mary B.
APPLICANT: Neuhaus, Jean-Marc
APPLICANT: Payne, George B.
APPLICANT: Sperison, Christoph
APPLICANT: Stinson, Jeffrey R.
APPLICANT: Uknes, Scott J.
APPLICANT: Ward, Eric R.
TITLE OF INVENTION: CHEMICALLY REGULATABLE AND ANTI-PATHOGENIC
TITLE OF INVENTION: DNA SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 106
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,043
FILING DATE: 24-MAY-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/181,271
FILING DATE: 13-JAN-94
APPLICATION NUMBER: US 08/093,301
FILING DATE: 16-JUL-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/937,197
FILING DATE: 6-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/678,378
FILING DATE: 1-APR-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/305,566
FILING DATE: 6-FEB-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/165,667
FILING DATE: 8-MAR-1988
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/042,847
FILING DATE: 6-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/632,441
FILING DATE: 21-DEC-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/425,504
FILING DATE: 20-OCT-1989
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/848,506
FILING DATE: 6-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/768,122
FILING DATE: 27-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/580,431
FILING DATE: 7-SEP-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/368,672
FILING DATE: 20-JUN-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/329,018
FILING DATE: 24-MAR-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/045,957
FILING DATE: 12-APR-1993
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: S-19825/PI/CGC 1727
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ. ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 4483 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-449-043-5

Query Match 76.0%; Score 15.2; DB 1; Length 4483;
Best Local Similarity 85.0%; Pred. No. 1.6e+02;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTATTAGT 20
||||| |||||
Db 1627 TTAGTGTAGTATTAGT 1646

Search completed: April 6, 2003, 11:56:32
Job time : 50.25 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 6, 2003, 11:24:32 ; Search time 58 Seconds
(without alignments)
302.472 Million cell updates/sec

Title: US-09-980-265-12

Perfect score: 20
Sequence: 1 ttagttagttagttagtgc 20

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 593429 seqs, 438583890 residues

Total number of hits satisfying chosen parameters: 1186858

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2-6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2-6/ptodata/1/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2-6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2-6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2-6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2-6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
7: /cgn2-6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2-6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2-6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
10: /cgn2-6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
11: /cgn2-6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
12: /cgn2-6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2-6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
14: /cgn2-6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	16.8	84.0	407	US-09-878-574-3626	Sequence 3626, App
2	16.8	84.0	513509	US-09-754-853A-4	Sequence 4, Appli
3	16.4	82.0	3376	US-09-834-975-775	Sequence 775, App
4	15.8	79.0	551	US-08-781-986A-994	Sequence 994, App
5	15.8	79.0	605	US-09-920-300A-827	Sequence 827, App
6	15.8	79.0	605	US-10-033-528-827	Sequence 827, App
7	15.8	79.0	657	US-09-938-842A-3055	Sequence 3055, App
8	15.8	79.0	1305	US-09-887-576-491	Sequence 491, App
9	15.8	79.0	1305	US-09-887-576-491	Sequence 549, App
10	15.8	79.0	2000	US-09-887-576-491	Sequence 97, Appl
11	15.8	79.0	35414	US-09-836-607-45	Sequence 45, Appl
12	15.4	77.0	380	US-09-998-598-2130	Sequence 2130, App
13	15.4	77.0	467	US-09-864-761-2265	Sequence 2265, App
14	15.2	76.0	323	US-09-960-352-11874	Sequence 11874, A
15	15.2	76.0	441	US-09-964-824A-191	Sequence 191, App
16	15.2	76.0	456	US-09-791-279-40	Sequence 40, Appl
17	15.2	76.0	548	US-09-924-035A-396	Sequence 396, App
18	15.2	76.0	646	US-10-033-528-1865	Sequence 1865, App
19	15.2	76.0	1022	US-08-781-986A-910	Sequence 910, App

20	15.2	76.0	1202	US-09-822-830A-591	Sequence 591, App
21	15.2	76.0	1557	US-10-001-873-1	Sequence 1, Appli
22	15.2	76.0	2000	US-09-938-842A-3169	Sequence 3169, Ap
23	15.2	76.0	2000	US-09-938-842A-3416	Sequence 3416, Ap
24	15.2	76.0	2000	US-09-938-842A-4038	Sequence 4038, App
25	15.2	76.0	2315	US-09-925-302-148	Sequence 148, App
26	15.2	76.0	6285	US-09-070-927A-415	Sequence 415, App
27	15.2	76.0	21393	US-10-091-526-15	Sequence 15, Appl
28	15.2	76.0	32193	US-09-764-877-2623	Sequence 2623, Ap
29	15.2	76.0	32463	US-09-996-956-5	Sequence 5, Appli
30	15.2	76.0	99916	US-09-816-095-3	Sequence 3, Appli
31	15.2	76.0	335913	US-09-754-853A-2	Sequence 2, Appli
32	15.2	76.0	335913	US-09-754-853A-3	Sequence 3, Appli
33	15.2	76.0	397658	US-09-813-320-3	Sequence 3, Appli
34	14.8	74.0	123	US-09-867-701-2222	Sequence 2222, Ap
35	14.8	74.0	177	US-09-867-701-3506	Sequence 3506, Ap
36	14.8	74.0	180	US-09-834-975-467	Sequence 4780, Ap
37	14.8	74.0	193	US-09-834-975-467	Sequence 467, App
38	14.8	74.0	203	US-10-012-896-717	Sequence 717, App
39	14.8	74.0	203	US-09-895-793-717	Sequence 717, App
40	14.8	74.0	203	US-09-895-793-717	Sequence 717, App
41	14.8	74.0	203	US-09-789-143-717	Sequence 717, App
42	14.8	74.0	203	US-09-780-669-717	Sequence 717, App
43	14.8	74.0	203	US-09-822-827-717	Sequence 717, App
44	14.8	74.0	207	US-09-834-975-338	Sequence 338, App
45	14.8	74.0	234	US-09-796-692-8941	Sequence 8941, Ap

ALIGNMENTS

RESULT 1
US-09-878-574-3626
Sequence 3626, Application US/09878574
Patent No. US20020110548A1

GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Thompson, Michael D.
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants
FILE REFERENCE: 38-21(15401)B
CURRENT APPLICATION NUMBER: US/09/878,574
CURRENT FILING DATE: 2001-12-21
PRIOR APPLICATION NUMBER: 09/333,535
PRIOR FILING DATE: 1999-06-14
NUMBER OF SEQ ID NOS: 15775
SEQ ID NO 3626
LENGTH: 407
TYPE: DNA
ORGANISM: Glycine max
OTHER INFORMATION: Clone ID: LIB3028-007-Q1-B1-H12
US-09-878-574-3626

Query Match 84.0%; Score 16.8; DB 10; Length 407;
Best Local Similarity 90.0%; Pred. No. 2.7e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGTT 20
11 |||||
DB 216 TTCTTTGTTAGTTATTAGTT 235

RESULT 2
US-09-754-853A-4/c
Sequence 4, Application US/09754853A
Publication No. US20030005491A1

GENERAL INFORMATION:
APPLICANT: Hauge, Brian M.
APPLICANT: Parnell, Laurence D.
APPLICANT: Parsons, Jeremy D.
APPLICANT: Wang, Ming Li
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With

;; TITLE OF INVENTION: Soybean Cyst Nematode Resistance
;; FILE REFERENCE: 38-10(15810)B
;; CURRENT APPLICATION NUMBER: US/09/754, 853A
;; CURRENT FILING DATE: 2001-01-05
;; PRIOR APPLICATION NUMBER: US 60/7174, 880
;; PRIOR FILING DATE: 2000-01-07
;; NUMBER OF SEQ ID NOS: 1119
;; SEQ ID NO 4
;; LENGTH: 513509
;; TYPE: DNA
;; ORGANISM: Glycine max
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (111805)..(113968)..(114684)..(115204)
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (1)..(513509)
;; OTHER INFORMATION: unsure at all n locations
;; OTHER INFORMATION: Clone ID: 318013_region_A3
US-09-754-853A-4

Query Match 84.0%; Score 16.8; DB 9; Length 513509;
Best Local Similarity 90.0%; Pred. No. 4.5e+02;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTTATTAGTT 20
Db 133170 TTAGTGTAGTTATTAGTT 133151

RESULT 3
US-09-834-975-775/C
;; Sequence 775, Application US/09834975
;; Patent No. US20020110815A1
;; GENERAL INFORMATION:
;; APPLICANT: Lillie, James
;; APPLICANT: Brown, Jeffrey
;; APPLICANT: Bolt, Andrew
;; APPLICANT: Van Hufel, Christophe
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS AND METHODS
;; TITLE OF INVENTION: FOR THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
;; TITLE OF INVENTION: OF HUMAN CANCERS
;; FILE REFERENCE: MRI-016B
;; CURRENT APPLICATION NUMBER: US/09/834, 975
;; CURRENT FILING DATE: 2001-04-13
;; PRIOR APPLICATION NUMBER: 60/197, 538
;; PRIOR FILING DATE: 2000-04-14
;; NUMBER OF SEQ ID NOS: 1046
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 775
;; LENGTH: 3376
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (1)...(3376)
;; OTHER INFORMATION: n = A,T,C or G
US-09-834-975-775

Query Match 82.0%; Score 16.4; DB 10; Length 3376;
Best Local Similarity 94.4%; Pred. No. 4.8e+02;
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 AGTTGTAGTTATTAGTT 20
Db 3261 AGTTGTAGTTATTAGTT 3244

RESULT 4
US-08-781-986A-994/C
;; Sequence 994, Application US/08781986A
;; Publication No. US20030054436A1
;; GENERAL INFORMATION:

;; APPLICANT: Charles Kunsch
;; TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
;; NUMBER OF SEQUENCES: 5255
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Human Genome Sciences, Inc.
;; STREET: 9410 Key West Avenue
;; CITY: Rockville
;; STATE: Maryland
;; COUNTRY: USA
;; ZIP: 20850
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
;; COMPUTER: HP Vectra 486/33
;; OPERATING SYSTEM: MSDOS version 6.2
;; SOFTWARE: ASCII text
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/781, 986A
;; FILING DATE:
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Benson, Bob
;; REGISTRATION NUMBER: 30,446
;; REFERENCE/DOCKET NUMBER: PB248PP
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (301) 309-8504
;; TELEFAX: (301) 309-8512
;; INFORMATION FOR SEQ ID NO: 994:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 551 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
US-08-781-986A-994

Query Match 79.0%; Score 15.8; DB 7; Length 551;
Best Local Similarity 89.5%; Pred. No. 7.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TTAGTGTAGTTATTAGT 19
Db 476 TTAGTGTAGTTATTAGT 458

RESULT 5
US-09-920-300A-827/C
;; Sequence 827, Application US/09920300A
;; Patent No. US20020136728A1
;; GENERAL INFORMATION:
;; APPLICANT: King, Gordon E.
;; APPLICANT: Meagher, Madeline Joy
;; APPLICANT: Xu, Jiangchun
;; APPLICANT: Secrist, Heather
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
;; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
;; FILE REFERENCE: 210121.547
;; CURRENT APPLICATION NUMBER: US/09/920, 300A
;; CURRENT FILING DATE: 2001-07-31
;; NUMBER OF SEQ ID NOS: 1789
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 827
;; LENGTH: 605
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-920-300A-827

Query Match 79.0%; Score 15.8; DB 10; Length 605;
Best Local Similarity 89.5%; Pred. No. 7.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 TTAGTGTAGTTATTAGTT 20

Db 578 TAGTAGTATTATCGTT 560

RESULT 6
US-10-033-528-827/c
; Sequence 827, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon F.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.547c1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 827
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-033-528-827

Query Match 79.0%; Score 15.8; DB 12: Length 605;
Best Local Similarity 89.5%; Pred. No. 7.1e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 TAGTTGTTAGTTATTAGTT 20
Db 578 TAGTAGTATTATCGTT 560

RESULT 7
US-09-938-842A-3055
; Sequence 3055, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kieps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: S1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 3055
; LENGTH: 657
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-3055

Query Match 79.0%; Score 15.8; DB 9: Length 657;
Best Local Similarity 89.5%; Pred. No. 7.2e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGT 19
Db 82 TTAGTTATTACTTATTAGT 100

RESULT 8
US-09-887-576-491

; Sequence 491, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.0010S1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 491
; LENGTH: 1305
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-491

Query Match 79.0%; Score 15.8; DB 10: Length 1305;
Best Local Similarity 89.5%; Pred. No. 7.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGT 19
Db 753 TTAGTTATTACTTATTAGT 771

RESULT 9
US-09-887-576-549
; Sequence 549, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.0010S1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 549
; LENGTH: 1305
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-549

Query Match 79.0%; Score 15.8; DB 10: Length 1305;
Best Local Similarity 89.5%; Pred. No. 7.7e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGT 19
Db 753 TTAGTTATTACTTATTAGT 771

Db 753 TTAGTTACTTATTAGT 771

RESULT 10

US-09-887-576-97
; Sequence 97, Application US/09887576
; Patent No. US20020144047A1
; GENERAL INFORMATION:
; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, Bret
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001us1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 97
; LENGTH: 2000
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-97

Query Match 79.0%; Score 15.8; DB 10; Length 2000;
Best Local Similarity 89.5%; Pred. No. 8e+02;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGT 19
Db 1426 TTAGTTACTTATTAGT 1444

RESULT 11

US-09-836-607-45/c
; Sequence 45, Application US/09836607
; Patent No. US20020098541A1
; GENERAL INFORMATION:
; APPLICANT: Ni, Jian
; APPLICANT: Ruben, Steven M.
; TITLE OF INVENTION: TNER Related Gene 12
; FILE REFERENCE: PF490P1
; CURRENT APPLICATION NUMBER: US/09/836,607
; CURRENT FILING DATE: 2001-04-17
; PRIOR APPLICATION NUMBER: 60/158,388
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: 09/421,112
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/104,950
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 45
; LENGTH: 35414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-836-607-45

Query Match 79.0%; Score 15.8; DB 10; Length 35414;
Best Local Similarity 89.5%; Pred. No. 1e+03;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TTAGTTGTTAGTTATTAGT 19
Db 1426 TTAGTTACTTATTAGT 1444

Db 16222 TTATTTGTTAGTTATTGCT 16204

RESULT 12

US-09-998-598-2130/c
; Sequence 2130, Application US/09998598
; Patent No. US20020150922A1
; GENERAL INFORMATION:
; APPLICANT: Stolk, John A.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Cheneault, Ruth A.
; APPLICANT: Meagher, Madelein Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.561
; CURRENT APPLICATION NUMBER: US/09/998,598
; CURRENT FILING DATE: 2001-11-16
; NUMBER OF SEQ ID NOS: 2606
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 2130
; LENGTH: 380
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-998-598-2130

Query Match 77.0%; Score 15.4; DB 10; Length 380;
Best Local Similarity 94.1%; Pred. No. 9.9e+02;
Matches 16; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 GTTGTAGTTATTAGTT 20
Db 93 GTTGTAGTCATTAGTT 77

RESULT 13

US-09-864-761-2265
; Sequence 2265, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeonica X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 2265
; LENGTH: 467
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL03383.24
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 6.4
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.4
US-09-864-761-2265
```

```
Query Match 77.0%; Score 15.4; DB 10; Length 467;
Best Local Similarity 94.1%; Pred. No. 1e+03; Indels 0; Gaps 0;
Matches 16; Conservative 0; Mismatches 1;
```

```
OY 3 AGTGTAGTATTAGT 19
   |||||||
DB 205 AGCTGTAGTATTAGT 221
```

```
RESULT 14
US-09-960-352-11874/c
; Sequence 11874, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 11874
; LENGTH: 323
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 51-LIB2809-009-Q1-E1-E4
US-09-960-352-11874
```

```
Query Match 76.0%; Score 15.2; DB 10; Length 323;
Best Local Similarity 85.0%; Pred. No. 1.2e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 TTAGTTGTTAGTTATTAGT 20
   |||||||
DB 88 TTGTTGTTAGTTAGGAGT 69
```

```
RESULT 15
US-09-964-824A-191/c
; Sequence 191, Application US/09964824A
; Patent No. US20020102531A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Horigan, Stephen
; TITLE OF INVENTION: Cancer Gene Determination and Therapeutic Screening Using Sign
; FILE REFERENCE: 689290-73
; CURRENT APPLICATION NUMBER: US/09/964,824A
; CURRENT FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: US/60/236,033
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,032
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US/60/236,028
; PRIOR FILING DATE: 2000-09-28
; NUMBER OF SEQ ID NOS: 583
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 191
; LENGTH: 441
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-964-824A-191
```

```
Query Match 76.0%; Score 15.2; DB 10; Length 441;
Best Local Similarity 85.0%; Pred. No. 1.2e+03;
Matches 17; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 1 TTAGTTGTTAGTTATTAGT 20
   |||||||
DB 242 TTAGTTGTTGCTTATTATT 223
```

```
Search completed: April 6, 2003, 12:14:20
Job time : 167 secs
```

THIS PAGE BLANK (USPTO)